



⑫

⑬ Anmeldenummer: 93116058.4

⑭ Anmeldetag: 05.10.93

EUROPÄISCHE PATENTANMELDUNG

⑮ Int. Cl. 5: C12N 7/00, C12N 15/48,
C07K 15/00, G01N 33/569,
C12Q 1/68, G01N 33/68

Der Anmelder hat nachträglich ein
Sequenzprotokoll eingereicht und erklärt, dass
dieses keine neuen Angaben enthält.

⑯ Priorität: 06.10.92 DE 4233848
22.10.92 DE 4235718
30.12.92 DE 4244541
01.08.93 DE 4318186

⑰ Veröffentlichungstag der Anmeldung:
13.04.94 Patentblatt 94/15

⑱ Benannte Vertragsstaaten:
AT BE CH DE DK ES FR GB IT LI LU NL PT SE

⑲ Anmelder: BEHRINGWERKE
Aktiengesellschaft
Postfach 1140
D-35001 Marburg(DE)

⑳ Erfinder: Gürtler, Lutz G., Prof. Dr.
Teufelsbergstrasse 18
D-81249 München(DE)
Erfinder: Eberle, Josef, Dr.
Sonnenstrasse 7c
D-85356 Freising(DE)
Erfinder: Brunn v., Albrecht, Dr.
Schumannstrasse 17
D-86154 Augsburg(DE)
Erfinder: Knapp, Stefan, Dr.
Wehrhäuser Strasse 6
D-35041 Marburg-Wehrhausen(DE)
Erfinder: Hauser, Hans-Peter, Dr.
Wannkopfstrasse 12
D-35037 Marburg(DE)

㉑ Vertreter: Keller, Günter, Dr et al
Lederer, Keller & Riederer
Patentanwälte
Prinzregentenstrasse 18
D-80538 München (DE)

㉒ Retrovirus aus der HIV-Gruppe und dessen Verwendung.

㉓ Offenbart wird ein neues Immunschwächevirus mit der Bezeichnung MVP-5180/91, das bei der European Collection of Animal Cell Cultures (ECACC) unter der Nr. V 920 92 318 hinterlegt wurde. Weiterhin offenbart werden die daraus erhältlichen charakteristischen Antigene, die für den Nachweis von Antikörpern gegen Retrovirus, die mit Immunschwächeerkrankungen verbunden sind, eingesetzt werden können sowie die DNS- und Aminosäuresequenz des Virus.

EP 0 591 914 A2

Die vorliegende Erfindung betrifft ein neues Retrovirus aus der HIV-Gruppe sowie Varianten oder Teile davon, welche die wesentlichen Eigenschaften des Virus enthalten. Beschrieben wird ein Verfahren zur Züchtung des Retrovirus. Die Erfindung betrifft weiterhin die Gewinnung dieses Retrovirus sowie die Verwendung des Virus, seiner Teile oder Extrakte für medizinische Zwecke, für die Diagnostik und bei der Herstellung von Impfstoffen.

Retroviren, die zur sogenannten HIV-Gruppe gehören, führen bei damit infizierten Menschen zu Krankheitsscheinungen, die unter dem Sammelbegriff Immunschwäche bzw. AIDS (Acquired Immune Deficiency Syndrome) zusammengefaßt werden.

10 Agents für die überwiegende Mehrheit der AIDS (Acquired Immune Deficiency Syndrome)-Fälle darstellt. Ein 1983 aus einem Patienten isoliertes und charakterisiertes Retrovirus erhielt die Bezeichnung HIV-1 (Barre-Sinoussi, F. et al., *Science* 220, 868-871 [1983]). Eine Variante von HIV-1 wird in WO 88/02383 beschrieben.

15 Eine zweite Gruppe von Humanen Immunschwäche Viren wurde 1985 in Westafrika identifiziert (Clavel, F. et al., *Science* 233, 343-346 [1986]) und als Humanes Immunschwäche Virus Typ 2 (HIV-2) bezeichnet (EP-A-0 239 425). HIV-2 Retroviren unterscheiden sich deutlich von HIV-1, weisen jedoch auch eine Verwandtschaft zu Affen Immunschwäche Viren (SIV-2) auf. Wie HIV-1 führt auch HIV-2 zu einer AIDS-Symptomatik.

20 Eine weitere Variante eines Immunschwäche Retrovirus wird in der EP-A-0 345 375 beschrieben und dort als HIV-3 Retrovirus bezeichnet (ANT 70). Auch in *Lancet* Vol. 340, Sept. 1992, S. 681-682 wird die Isolierung eines weiteren, variablen Immunschwächevirus beschrieben.

25 Es ist ein Charakteristikum der Humanen Immunschwäche Viren, daß sie eine hohe Variabilität aufweisen, die die Vergleichbarkeit der verschiedenen Isolate deutlich kompliziert. Beim Vergleich diverser HIV-1-Isolate treten z.B. in einigen Regionen des Genoms hohe Variabilitäten auf, während andere Genombereiche vergleichsweise konserviert vorliegen (Benn, S. et al. *Science* 230, 949-951 [1985]). Ein wesentlich größerer Polymorphismus konnte auch für HIV-2 beobachtet werden (Clavel, F. et al., *Nature* 324, 691-695 [1986]). Die größte genetische Stabilität besitzen Bereiche in den gag und pol Genen, die für strukturell und enzymatisch essentielle Proteine codieren; einige Regionen im env-Gen sowie die Gene (vif, vpr, tat, rev, nef), die für regulatorische Proteine codieren, zeigen einen hohen Grad an Variabilität. Es konnte weiterhin gezeigt werden, daß Antisera gegen HIV-1 auch mit gag und pol Genprodukten von HIV-2 kreuzreagieren, obwohl nur geringe Sequenzhomologie vorlag. Ebenfalls war die Hybridisierung zwischen diesen beiden Viren wenig signifikant, wenn nicht sehr wenig stringenten Konditionen verwandt wurden (Clavel, F. et al., *Nature* 324, 691-695 [1986]).

35 Aufgrund der weiten Verbreitung der Retroviren aus der HIV-Gruppe und der Tatsache, daß zwischen dem Zeitpunkt der Infektion und dem Zeitpunkt, zu dem eindeutige Symptome für pathologische Veränderungen erkennbar sind, ein Zeitraum von einigen bis vielen Jahren (2-20) liegt, ist es epidemiologisch von großer Bedeutung, die Infektion mit Retroviren der HIV-Gruppe möglichst frühzeitig und vor allem zuverlässig zu bestimmen. Dies spielt nicht nur eine Rolle bei der Diagnose von Patienten, die Zeichen von Immunschwäche aufweisen, sondern auch bei der Überprüfung von Blutspendern. Es hat sich herausgestellt, daß bei der Verwendung von Retroviren oder Bestandteilen davon des Typs HIV-1 oder HIV-2 in Nachweissystemen bei manchen Seren kein oder nur ein schwacher Nachweis von Antikörpern geführt werden kann, obwohl bei den Patienten, von denen die Seren stammen, Zeichen von Immunschwäche auftreten. Mit Hilfe des erfindungsgemäßen Retrovirus aus der HIV-Gruppe ist in bestimmten Fällen ein derartiger Nachweis möglich.

40 Beschrieben wird die Isolation und Charakterisierung eines neuen Humanen Immunschwäche Virus, im folgenden als MVP-5180/91 bezeichnet, das aus peripheren Lymphozyten einer 1991 34-jährigen Patientin aus Kamerun isoliert wurde, die Zeichen von Immunschwäche aufwies. Geographisch stammt dieses Retrovirus aus einer Region in Afrika, die zwischen Westafrika mit endemischer HIV-2 und HIV-1 Virusinfektion und Ostzentralafrika mit fast ausschließlicher HIV-1-Verbreitung lokalisiert ist. Gegenstand der vorliegenden Erfindung ist also ein neues Retrovirus der HIV-Gruppe, welches als MVP-5180/91 bezeichnet wird und dessen Varianten sowie davon abgeleitete DNS-Sequenzen und Aminosäuresequenzen bzw. Teilessequenzen und diese enthaltende Test-Kits. Das Retrovirus MVP-5180/91 wurde bei der European Collection of Animal Cell Cultures (ECACC) unter der Nummer V 920 92 318 gemäß den Bedingungen des

45 Budapest Vertrages hinterlegt. Ebenso wie HIV-1 und HIV-2 wächst das erfindungsgemäße MVP-5180/91 in folgenden Zelllinien HUT 78, Jurkat-Zellen, C8166-Zellen und MT-2-Zellen. Die Isolierung und Vermehrung von Viren wird in dem Buch "Viral Quantitation in HIV Infection, Editor Jean-Marie Andrieu, John Libbey Eurotext, 1991" ausführ-

lich beschrieben. Die dort beschriebenen Arbeitsmethoden werden durch Bezugnahme zum Gegenstand der Offenbarung der vorliegenden Anmeldung gemacht.

Weiterhin besitzt das erfindungsgemäße Virus eine magnesiumabhängige Reverse Transkriptase, die aber nicht manganabhängig ist. Dies stellt eine weitere Gemeinsamkeit zu den Viren HIV-1 und HIV-2 dar.

5 Zum besseren Verständnis der Unterschiede des erfindungsgemäßen MVP-5180/91 Virus zu den Retroviren HIV-1 und HIV-2 soll zunächst kurz der Aufbau der Immunschwäch verursachenden Retroviren erläutert werden. Im Inneren des Virus befindet sich die RNA in einem kugelförmigen Core, das aus Proteinuntereinheiten zusammengesetzt ist, die die Bezeichnung p 24 (p für Protein) tragen. Dieses innere Core wird von einer Proteinhülle umgeben, die aus dem Protein p 17 aufgebaut ist (äußeres Core) und von 10 einer Glykoproteinhülle umgeben ist, die neben Lipiden, die aus der Wirtszelle stammen, das Transmembranprotein gp 41, und das Hüllprotein 120 (gp 120) enthält. Dieses gp 120 kann dann mit den CD-4-Rezeptoren der Wirtszellen eine Bindung eingehen.

15 Soweit bekannt, weist die RNA der HIV-Viren - vereinfacht dargestellt - folgende Genbereiche auf: An den beiden Enden sogenannte long terminal repeats (LTR), und die folgenden Genbereiche gag, pol, env und nef. Das Gen gag codiert unter anderem für die Kern(Core)-Proteine, p 24 und p 17, das Gen pol codiert u.a. für die Reverse Transkriptase, die RNase H und die Integrase und das Gen env codiert für die Glykoproteine gp 41 und gp 120 der Virushülle. Das Gen nef codiert für ein Protein mit Regulatorfunktion. Eine schematische Anordnung des Genomes von Retroviren des HIV-Typs ist in Figur 1 gezeigt.

20 Eine Unterscheidung zwischen den Retroviren HIV-1 und HIV-2 ist u.a. dadurch möglich, daß virales Antigen ausgetestet wird mit einem monokonalen Antikörper, der kommerziell als-Testkit von Abbott (HIVAG-1 Monoclonal) erhältlich ist, und gegen das (HIV-1) p 24 gerichtet ist. Es ist bekannt, daß der Gehalt an Reverser Transkriptase in den Virustypen HIV-1 und HIV-2 etwa gleich ist. Wenn man deshalb in Verdünnungen der aufgeschlossenen Viren die Extinktion (E 490 nm), erhalten durch die Antigen-Antikörper-Reaktion, aufträgt gegen die Aktivität der Reversen Transkriptase, dann erhält man eine Graphik, die 25 etwa der Figur 2 entspricht. Hierbei stellt man fest, daß im Verhältnis zu dem Gehalt an Reverser Transkriptase bei HIV-1 eine sehr hohe Bindungsaffinität für p 24 mit dem eingesetzten monokonalen Antikörper vorhanden ist. Für HIV-2 tritt dagegen nur eine sehr geringe Bindungsaffinität für p 24 bei Einsatz des monokonalen Antikörpers wiederum bezogen auf den Gehalt an Reverser Transkriptase auf. Werden diese Messungen durchgeführt für MVP-5180/91, dann befindet sich die Kurve ziemlich genau in 30 der Mitte zwischen der Kurve von HIV-1 und HIV-2, d.h. die Bindungsaffinität des monokonalen Antikörpers gegen MVP-5180/91 p 24 ist gegenüber HIV-1 reduziert. Figur 2 zeigt schematisch diesen Sachverhalt, wobei RT Reverse Transkriptase bedeutet und als Antigen (Ag) das Protein p 24 eingesetzt wird, gegen das der monoklonale Antikörper, der in dem von Abbott käuflich erwerblichen Testkit vorhanden ist, gerichtet ist.

35 Ein sehr vielseitig verwendbares System der Gentechnologie ist die sogenannte PCR (polymerase chain reaction) geworden, wobei die zur Durchführung des Verfahrens benötigten Komponenten käuflich erworben werden können. Mit diesem Verfahren ist es möglich, DNA-Sequenzen zu amplifizieren, wenn DNA-Bereiche der zu amplifizierenden Sequenz bekannt sind. Es müssen dann kurze komplementäre DNA-Fragmente (Oligonukleotide = Primer) synthetisiert werden, die sich an einen kurzen Bereich der zu amplifizierenden Nukleinsäuresequenz anlagern. Für die Testdurchführung werden HIV-Nukleinsäuren mit 40 den Primern zusammengebracht in einer Reaktionsmischung, die zusätzlich eine Polymerase und Nukleotidtriphosphate enthält. Die Polymerisation (DNA-Synthese) wird für eine bestimmte Zeit durchgeführt und dann werden die Nukleinsäurestränge durch Erwärmen getrennt. Nach Abkühlen läuft dann die Polymerisation erneut an. Wenn es sich also bei dem erfindungsgemäßen Retrovirus um ein HIV-1 oder HIV-2 Virus handelt, dann müßte die Nukleinsäure amplifiziert werden können, indem Primer verwendet werden, die 45 konserviert sind innerhalb der bekannten Sequenzen der Viren HIV-1 und HIV-2. Derartige Primer sind zum Teil vorbeschrieben (Lauré, F. et al., Lancet ii, (1988) 538-541 für pol 3 und pol 4 bzw. Ou C.Y. et al., Science 239 (1988) 295-297 für sk 38/39, sk 68/69).

Es wurde nun herausgefunden, daß bei Verwendung von bestimmten Primerpaaren, die folgende Sequenz aufweisen:

50

55

(Seq. ID No. 1-14)

HIV-1

5 gaga: CTACT AGTAC CCTTC AGG
 gagb: CGGTC TACAT AGTCT CTAAA G
 sk38: CCACC TATCC CAGTA GGAGA A
 10 sk39: CCTTT GGTCC TTGTC TTATG TCCAG AATGC oder

15 pol3: TGGGA AGTTC AATTA GGAAT ACCAC
 pol4: CCTAC ATAGA AATCA TCCAT GTATT G
 20 pol3n: TGGAT GTGGG TGATG CATA
 pol4n: AGCAC ATTGT ACTGA TATCT A sowie

25 SK145: AGTGG GGGGA CATCA AGCAG CC
 SK150: TGCTA TGTCA CTTCC CCTTG GT
 30 145-P: CCATG CAAAT GTTAA AAGAG AC
 150-P: GGCCT GGTGC AATAG GCCC

35 oder eine Kombination von pol3 und pol4 mit

40 UNI-1: GTGCT TCCAC AGGGA TGGAA
 UNI-2: ATCAT CCATG TATTG ATA

(Donehower L.A. et al. (1990) J. Virol. Methods 28, 33-46)
 und mit der PCR mit nested primer schwache Amplifikate der MVP-5180/91-DNA erhalten wurden.
 Keine oder nur schwache Amplifikate im Vergleich zum HIV-1, die evtl. auf Verunreinigungen zurückzu-
 führen sind, wurden erhalten mit folgenden Primer-Sequenzen:

50

55

(Seq. ID No. 15-34)

5 tat 1 AATGG AGCCA GTAGA TCCTA
 tat 2 TGTCT CCGCT TCTTC CTGCC

10 tat 1P GAGCC CTGGA AGCAT CCAGG
 tat 2P GGAGA TGCCT AAGGC TTTTG

15 enva: TGTTC CTTGG GTTCT TG
 envb: GAGTT TTCCA GAGCA ACCCC
 sk68: AGCAG CAGGA AGCAC TATGG
 sk69: GCCCC AGACT GTGAG TTGCA ACAG

20 5v3e: GCACA GTACA ATGTA CACAT GG
 3v3e: CAGTA GAAAA ATTCC CCTCC AC
 25 5v3degi: TCAGG ATCCA TGGGC AGTCT AGCAG AAGAA G
 3v3degi: ATGCT CGAGA ACTGC AGCAT CGATT CTGGG TCCCC TCCTG AG
 3v3longdegi: CGAGA ACTGC AGCAT CGATG CTGCT CCCAA GAACC CAAGG
 3v3longext: GGAGC TGCTT GATGC CCCAG A

30 gagdi: TGATG ACAGC ATGTC AGGGA GT
 pol e: GCTGA CATTT ATCAC AGCTG GCTAC

35 Im Vergleich zum HIV-1 schwache Amplifikate, die jedoch die gleiche Intensität wie das verwendete HIV-2 Isolat (MVP-11971/87) aufwiesen, wurden erhalten mit

40 gag c: TATCA CCTAG AACTT TAAAT GCATG GG
 gag d: AGTCC CTGAC ATGCT GTCAT CA
 env c: GTGGA GGGGA ATTTT TCTAC TG
 env d: CCTGC TGCTC CCAAG AACCC AAGG.

50 Eine weitverbreitete Methode zum Nachweis von HIV-Antikörpern ist der sogenannte Western Blot (Immunoblot). Dabei werden die viralen Proteine geelektrophoretisch aufgetrennt und dann auf eine Membran überführt. Die mit den überführten Proteinen versehenen Membranen werden dann mit Sera der zu untersuchenden Patienten in Verbindung gebracht. Sofern Antikörper gegen die viralen Proteine vorhanden sind, binden diese an die Proteine. Nach Waschen verbleiben lediglich spezifische Antikörper gegen virale Proteine. Die Antikörper werden dann mit Antikörpern sichtbar gemacht, die regelmäßig mit einem Enzym gekoppelt sind, das eine Farbreaktion katalysiert. Auf diese Weise können die Banden der viralen Proteine sichtbar gemacht werden.

55 Das erfindungsgemäße Virus MVP-5180/91 weist gegenüber den Viren HIV-1 und HIV-2 im Western Blot zwei signifikante wesentliche Unterschiede auf. Das HIV-1 zeigt regelmäßig eine starke Bande, die dem Protein p 24 zuzuordnen ist und eine sehr schwache, oft kaum sichtbare Bande, die dem Protein p 23 zuzuordnen ist. HIV-2 weist eine kräftige Bande auf, die dem Protein p 25 zuzuordnen ist und manchmal

ein schwache Bande, die dem Protein p 23 zuzuordnen ist. Im Unterschied dazu weist das erfindungsgemäße MVP-5180/91 Virus zwei etwa gleich starke Banden auf, die den Proteinen p 24 und p 25 entsprechen.

Ein weiterer signifikanter Unterschied besteht bei den Banden, die der Reversen Transkriptase zuzuordnen sind. HIV-1 zeigt eine Bande (p 53), die der Reversen Transkriptase entspricht und eine Bande (p 66), die der Reversen Transkriptase verbunden mit der RNase H entspricht. Bei HIV-2 entspricht die Reverse Transkriptase dem Protein p 55 und, wenn sie mit der RNase H verbunden ist, dem Protein p 68. Das erfindungsgemäße MPV-5180/91 weist dagegen eine Bande auf bei dem Protein p 48, die der Reversen Transkriptase entspricht, und eine Bande, bei dem Protein p 60, die der Reversen Transkriptase in Verbindung mit RNase H entspricht.

Aus diesen Ergebnissen kann geschlossen werden, daß die Reverse Transkriptase des MVP-5180/91 ein Molekulargewicht hat, das zwischen etwa 3 und etwa 7 Kilodalton kleiner ist als das der Reversen Transkriptase von HIV-1 bzw. HIV-2. Die Reverse Transkriptase von MPV-5180 weist also ein Molekulargewicht auf, das zwischen etwa 4.500 Dalton und etwa 5.500 Dalton kleiner ist als die Reverse Transkriptase von HIV-1 bzw. HIV-2.

Es wurde herausgefunden, daß mit Hilfe des erfindungsgemäßen Virus MVP-5180/91 anti-env-Antikörper in Seren von deutschen Patienten, die Zeichen von Immunschwäche zeigen, nur schwach nachgewiesen werden können, wobei die Seren aber stark reagieren, wenn anstelle des erfindungsgemäßen Virus ein HIV-1 Virus verwendet wird. Diese stärkere Nachweisreaktion wurde vor allem lokalisiert in dem gp 41 Protein. Bei den Versuchen wurden Serumpools gegenübergestellt, die einmal von deutschen Patienten stammen und zum anderen von afrikanischen Patienten mit Zeichen von Immunschwäche stammen.

Die oben angegebenen Charakteristika kennzeichnen solche Virus-Varianten, die dem erfindungsgemäßen MVP-5180/91 entsprechen. Wenn also aus heparinisiertem Spenderblut, das von Personen stammt, die Immunschwächezeichen aufweisen und vorzugsweise aus Afrika stammen, Immunschwächeviren isoliert werden, dann kann auf diese Weise das erfindungsgemäße Virus oder Varianten davon erhalten werden.

Da das Virus isoliert wurde, welches die oben erwähnten Eigenschaften aufweist, kann die Klonierung einer cDNA auf folgendem Weg durchgeführt werden: Das Virus wird aus einer entsprechend großen Kulturmenge (etwa 1 l) präzipitiert und in phosphatgepufferter Kochsalzlösung aufgenommen. Dann erfolgt eine Pelletierung durch ein (20 %iges) Saccharose-Kissen. Das Viruspellet kann in 6 M Guanidiniumchlorid in 20 mM Dithiothreitol und 0,5 % Nonidet P 40 suspendiert werden. CsCl wird bis auf eine Konzentration von 2 molar zugegeben und die das aufgebrochene Virus enthaltende Lösung wird auf ein Cäsiumchlorid-Kissen aufgebracht. Dann wird die virale RNA durch Zentrifugation pelletiert, gelöst, mit Phenol extrahiert und mit Ethanol und Lithiumchlorid präzipitiert. Mit Hilfe eines Oligo(dT)-Primers wird die Synthese des ersten cDNA-Stranges an der viralen RNA oder Teilen davon durchgeführt. Die Synthese unter Zugabe von Reverser Transkriptase kann durchgeführt werden unter Verwendung eines käuflich erhältlichen Kits. Für die Synthese des zweiten Stranges wird der RNA-Strang des RNA/DNA-Hybrids mit RNase H verdaut und der zweite Strang unter Einsatz von E.coli DNA Polymerase I synthetisiert. Mit Hilfe von T4 DNA Polymerase können dann stumpfe Enden erzeugt werden und diese mit geeigneten Linkern für Restriktions-schnittstellen verbunden werden. Nach Restriktionsverdau mit der geeigneten Restriktionsendonuklease wird das cDNA-Fragment aus einem Agarosegel isoliert und mit einem vorher in geeigneter Weise geschnittenen Vektor ligiert. Der Vektor mit dem cDNA-Insert kann dann zur Transformation von kompetenten E.coli-Zellen verwendet werden. Die erhaltenen Kolonien werden dann auf Membranen übertragen, lysiert und denaturiert und schließlich durch Hybridisierung mit Digoxigenin oder Biotin markierter Nukleinsäure aufgefunden. Nach gentechnologischer Herstellung der entsprechenden cDNA ist eine Isolierung der gewünschten DNA-Fragmente, die aus dem Retrovirus stammen, möglich. Durch Einbau dieser Fragmente in geeignete Expressionsvektoren kann dann das gewünschte Protein bzw. Proteinfragment exprimiert werden und für die diagnostischen Tests eingesetzt werden.

Alternativ zu der angegebenen Methode kann das Immunschwächevirus mit Hilfe der PCR-Technologie kloniert werden, wobei die oben angegebenen Primer Verwendung finden können.

Zwischen verschiedenen Virusisolaten kann die Ähnlichkeit ausgedrückt werden durch den Grad der Homologie der Nucleinsäure- oder Proteinsequenzen. Eine 50 %ige Homologie bedeutet beispielsweise, daß 50 von 100 Nucleotid- oder Aminosäure-Positionen in den Sequenzen übereinstimmen. Die Homologie von Proteinen wird durch Sequenzanalyse bestimmt. Homologe DNA-Sequenzen lassen sich auch durch die Hybridisierungs-Technik ermitteln.

Erfindungsgemäß wurde zunächst ein Teil aus dem Hüllprotein sequenziert und festgestellt, daß diese Sequenz nur eine verhältnismäßig geringe Homologie zu den entsprechenden Sequenzen von Viren vom HIV-Typ aufweist. Insbesondere bezogen auf den gp 41-Bereich wurde durch einen Vergleich mit HIV-Sequenzen, der mit Hilfe von Datenbanken durchgeführt wurde, eine Homologie von höchstens 66 %

(Nucleotidsequenz) ermittelt.

Es wurde weiterhin der Bereich sequenziert, der für gp 41 kodiert. Diese Sequenz ist in Tabellen 1 bzw. 3 dargestellt.

Gegenstand der vorliegenden Erfindung sind daher solche Viren, die eine Homologie von mehr als 66 %, bevorzugt 75 % und besonders bevorzugt 85 % aufweisen zu dem erfindungsgemäßen HIV-Virus, MVP 5180/91, bezogen auf die Nucleotidsequenz der Tabelle 1 und/oder der Tabelle 3.

Weiterhin sind Gegenstand der vorliegenden Erfindung solche Viren, die eine Homologie von mehr als 66 %, bevorzugt 75 % und besonders bevorzugt 85 % aufweisen zu Partialsequenzen der in Tabelle 3 dargestellten Nucleotidsequenz, die wenigstens 50, bevorzugt 100 Nucleotide lang sind. Dies entspricht 10 einer Länge der Peptide von wenigstens 18 und bevorzugt von wenigstens 33 Aminosäuren.

Das erfindungsgemäße Virus unterscheidet sich durch seine Sequenz von vorbekannten Viren. Gegenstand der vorliegenden Erfindung sind daher solche Viren sowie entsprechende DNS- bzw. Aminosäuresequenzen, die der Sequenz des erfindungsgemäßen Virus weitgehend entsprechen, wobei der Grad der Abweichung durch den Grad der Homologie festgelegt wird. Eine Homologie von beispielsweise mehr als 15 85 % bedeutet daher, daß solche Sequenzen umfaßt werden, die in wenigstens 85 von 100 Nucleotiden bzw. Aminosäuren dieselben Nucleotide bzw. Aminosäuren aufweisen, während der Rest unterschiedlich sein kann. Bei der Feststellung der Homologie werden die beiden Sequenzen derart gegenübergestellt, daß möglichst viele einander entsprechende Nucleotide bzw. Aminosäuren miteinander zur Deckung kommen.

Die (nahezu) vollständige Sequenz, angegeben als DNS-Sequenz des erfindungsgemäßen Virus ist in 20 Fig. 4 wiedergegeben. Gegenstand der vorliegenden Erfindung sind dabei Viren, die die Sequenz gemäß Fig. 4 aufweisen sowie Varianten davon, die eine hohe Homologie zu der Sequenz von Fig. 4 aufweisen sowie davon abgeleitete Proteine, Polypeptide und Oligopeptide, die diagnostisch verwendbar oder als Impfstoff einsetzbar sind.

Anhand der isolierten Sequenz können immundominante Epitope (Peptide) konfektioniert und synthetisiert werden. Da die Nucleinsäuresequenz des Virus bekannt ist, kann der Fachmann hieraus die Aminosäuresequenz ableiten. Ein Teilbereich der Aminosäuresequenz ist in Tabelle 3 angegeben. Gegenstand der vorliegenden Erfindung sind daher auch Antigene, d.h. Proteine, Oligo- oder Polypeptide, die mit Hilfe der in Figur 4 bzw. Tabelle 3 offenbarten Information hergestellt werden können. Diese Antigene, Proteine, Polypeptide und Oligopeptide weisen Aminosäuresequenzen auf, die von Figur 4 abgeleitet werden können 30 bzw. in Tabelle 3 angegeben sind. Die Antigene bzw. Peptide können verhältnismäßig kurze Teilsequenzen einer Aminosäuresequenz aufweisen, die in Tabelle 3 wiedergegeben ist oder aus Figur 4 abgeleitet werden kann. Diese Aminosäuresequenz ist wenigstens 6 Aminosäuren, bevorzugt wenigstens 10 und besonders bevorzugt wenigstens 15 Aminosäuren lang. Hergestellt werden können diese Peptide nicht nur mit Hilfe der rekombinanten Technologie sondern auch durch synthetische Methoden. Ein geeigneter Herstellungs- 35 weg ist die Festphasensynthese vom Merrifield-Typ. Eine weitere Beschreibung dieser Technik und anderer im Stand der Technik bekannter Verfahren kann in der Literatur gefunden werden, z.B. M. Bodansky, et al., Peptide Synthesis, John Wiley & Sons, 2nd Edition 1976.

Bei den diagnostischen Tests wird eine Serumprobe der zu untersuchenden Person zusammengebracht mit den Proteinketten von einem oder mehreren Proteinen oder Glykoproteinen (die in eukaryontischen 40 Zelllinien exprimiert werden können) oder Teilen davon, die von MVP-5180/91 stammen. Bevorzugte Testverfahren schließen die Immunfluoreszenz oder immunenzymatische Testverfahren (z.B. Elisa, Immuno-blot) ein.

Bei den immunenzymatischen Tests (ELISA) kann beispielsweise Antigen, das von MVP-5180/91 oder einer Variante davon stammt, an den Wänden von Mikrotiterplatten gebunden werden. Die dabei verwendete Dosisierung hängt von dem Testsystem und der Behandlung der Mikrotiterplatten wesentlich ab. Dann wird Serum bzw. Serumverdünnungen, die von der zu untersuchenden Person stammen, in die Löcher der Mikrotiterplatten gegeben. Nach einer bestimmten Inkubationszeit wird die Platte gewaschen und spezifische Immunkomplexe werden nachgewiesen durch Antikörper, die spezifisch an menschliche Immunglobuline binden und die vorher mit einem Enzym, beispielsweise Meerrettichperoxidase, alkalischer Phosphatase usw. verbunden wurden oder mit enzymmarkiertem Antigen. Diese Enzyme können ein farbloses Substrat in ein stark gefärbtes Produkt umwandeln und an der Stärke der Färbung kann dann das Vorhandensein von spezifischen Anti-HIV-Antikörpern abgelesen werden. Eine weitere Möglichkeit der Verwendung des erfindungsgemäßen Virus in Testsystemen ist die Verwendung in Western Blots.

Auch wenn die Herstellung von Impfstoffen gegen Immunschwächeerkrankungen sich als äußerst schwierig erweist, kann doch auch dieses Virus bzw. Teile davon, d.h. immundominante Epitope und Induktoren der zellulären Immunität, oder gentechnologisch hergestellte Antigene zur Entwicklung und Herstellung von Impfstoffen verwendet werden.

B Beispiel 1

Das erfindungsgemäße Immunschwäche-Virus MVP-5180/91 wurde aus dem Blut einer Patientin mit Zeichen von Immunschwäche isoliert. Hierzu wurden periphera mononukleäre Zellen (peripheral blood lymphocytes, PBL) und periphera Lymphozyten aus dem Blut (PBL) eines nicht HIV-infizierten Spenders mit Phytohämagglutinin stimuliert und in Kultur gehalten. Verwendet wurde hierzu das üblich Medium RPMI 1640 mit 10 % fötalem Kälberserum. Die Kulturbedingungen sind beschrieben in Landay A. et al., J. Inf. Dis., 161 (1990) S. 706-710. Beobachtet wurde dann die Bildung von Riesenzellen unter dem Mikroskop. Die Produktion von HIV-Viren wurde über die Bestimmung des p 24-Antigens mit Hilfe des käuflich erwerbaren Tests von Abbott bestimmt. Ein weiterer Test zur Bestimmung des Wachstums der Viren war der Test unter Verwendung von partikelgebundener Reverser Transkriptase (Eberle J., Seibl R., J. Virol. Methods 40, 1992, S. 347-358). Das Wachstum der Viren wurde also anhand der enzymatischen Aktivitäten im Kulturüberstand ein- bis zweimal pro Woche bestimmt, um die Virusproduktion zu überwachen. Wöchentlich einmal wurden neue Spenderlymphozyten zugegeben.

15 Nachdem eine HIV-Virenvermehrung festgestellt werden konnte, wurden frische periphera Lymphozyten aus dem Blut (PBL) nicht HIV-infizierter, gesunder Spender mit dem Überstand der ersten Kultur infiziert. Dieser Schritt wurde wiederholt und dann wurden mit dem Überstand H 9 bzw. HUT 78 Zellen infiziert. Auf diese Weise war eine permanente Produktion des Immunschwäche-Virus möglich. Das Virus wurde bei der ECACC unter der Nr. V 920 92 318 hinterlegt.

20

Beispiel 2

Zum Nachweis von HIV-Infektionen ist derzeit der sogenannte Western Blot oder Immunoblot ein Standardverfahren. Gemäß der in J. Virol. Meth. 15 (1987) S. 11-23 von Gütler et al. beschriebenen Vorgehensweise wurden verschiedene Seren untersucht. Hierbei wurden Seren von deutschen Patienten solchen Seren gegenübergestellt, die von afrikanischen Patienten erhalten wurden. Es wurden hierbei folgende Ergebnisse erhalten:

Virustyp	deutsche Seren	afrikanische Seren
HIV-1, Virus isoliert von deutschem Patienten MVP-5180/91	starke Reaktion	starke Reaktion mit gp 41
	keine bis schwache Reaktion mit gp 41	starke Reaktion

30 Die oben dargestellten Ergebnisse zeigen, daß ein aus deutschen Patienten isoliertes Virus vom HIV-1 Typ, wenn es zum Nachweis von HIV-Infektionen verwendet wird, möglicherweise keine eindeutigen Ergebnisse liefert, wenn der Patient infiziert wurde mit einem Virus, das dem erfindungsgemäßen MVP-5180/91 entspricht. Es wird dabei davon ausgegangen, daß mit Hilfe des erfindungsgemäßen Virus solche Viren nachgewiesen werden können, die wenigstens etwa 85 % Homologie, bezogen auf das Gesamtgenom, zu dem erfindungsgemäßen Virus aufweisen.

Beispiel 3

45 Gemäß der in Beispiel 2 angegebenen Vorgehensweise wurden weitere Western Blots durchgeführt. Die Ergebnisse sind in der anliegenden Figur 3 dargestellt. Bei diesem Test wurde einmal das virale Protein des erfindungsgemäßen Immunschwäche-Virus MVP-5180/91 und einmal das virale Protein eines HIV-1-Typ Virus (MVP-899) geelektrophoretisch aufgetrennt und dann auf Zellulosefilter transferiert. Diese Filterstreifen wurden mit den Seren von verschiedenen Patienten inkubiert und dann wurden die spezifischen Antikörper durch eine Farbreaktion sichtbar gemacht. Die linke Hälfte der Figur mit der Überschrift MVP-5180 zeigt das erfindungsgemäße Immunschwäche-Virus. Die rechte Hälfte der Figur zeigt ein aus deutschem Spender isoliertes Virus (MVP-899), bei dem es sich um ein HIV-1-Virus handelt.

50 Die einzelnen Filterstreifen wurden nun mit den Seren von verschiedenen Patienten inkubiert. Betrachtet man die Figur 3, so wurden jeweils dieselben Seren (von deutschen Patienten) umgesetzt mit je zwei Filterstreifen, wobei die Nummern 8 und 26; 9 und 27; 10 und 28; 11 und 29; 12 und 30; 13 und 31; 14 und 32; 15 und 33 sowie 16 und 34 die gleichen Seren bezeichnen. Bei den Western Blots mit den Nummern 17 und 18 wurden Seren von afrikanischen Patienten eingesetzt. Die Zahlen an den rechten Seitenrändern geben die annähernden Molekulargewichte in Tausend (KD) an.

Die Figur 3 zeigt deutlich, daß Seren von deutschen Patienten mit dem erfindungsgemäßen Immunschwäche-Virus im Western Blot mit dem gp 41 nur sehr schwach reagieren. Seren von afrikanischen Patienten dagegen reagieren mit dem erfindungsgemäßen Immunschwäche-Virus sehr stark. Figur 3 macht daher deutlich, daß unter Verwendung des erfindungsgemäßen Immunschwäche-Virus solche Immunschwäche-Infektionen nachgewiesen werden können, die bei Verwendung eines HIV-1 oder HIV-2-Virus nur fragliche, also nicht eindeutig positive Ergebnisse liefern. Diese Nachweismöglichkeit kann weitreichende diagnostische Bedeutung haben, da in den Fällen, in denen nur fragliche Ergebnisse im Western Blot erzielt werden, nicht mit eindeutiger Sicherheit festgestellt werden kann, ob es sich um eine Infektion mit einem Immunschwäche-Virus handelt. Wenn aber mit Hilfe des erfindungsgemäßen Immunschwäche-Virus derartige fragliche Ergebnisse einer Infektion mit einem Virus des erfindungsgemäßen Typs zugeordnet werden können, dann stellt dies einen erheblichen diagnostischen Fortschritt dar.

Beispiel 4

15 DNA-Isolierung, Amplifizierung und strukturelle Charakterisierung von Genomabschnitten des HIV-Isolates MVP-5180/91

Genomische DNA aus MVP-5180/91-infizierten HUT 78-Zellen wurden nach Standardmethoden isoliert. Zur Charakterisierung von Genombereichen des Isolates MVP-5180/91 wurden PCR (Polymerase Chain

20 Reaction)-Experimente mit einem Primerpaar aus dem Hüllproteinbereich gp 41 durchgeführt. Die Durchführung der PCR-Experimente erfolgte nach der Methode von Saiki et al. (Saiki et al., Science 239: 487-491, 1988) mit folgenden Modifikationen: Für die Amplifikation HIV-spezifischer DNA-Bereiche wurden 5 µl 25 genomische DNA aus MVP-5180/91 infizierten HUT 78-Zellen in einem 100 µl Reaktionsansatz (0,25 mM dNTP, je 1 µM Primer 1 und Primer 2, 10 mM Tris HCl pH 8,3, 50 mM KCl, 1,5 mM MgCl₂, 0,001 % Gelatine, 2,5 units Taq Polymerase (Perkin Elmer) pipettiert und nach folgendem Temperatur-Programm 30 amplifiziert: 1. Initiale Denaturierung: 3' 95°C, 2. Amplifikation: 90" 94°C, 60" 56°C, 90" 72°C (30 Cycles).

Die für die PCR und die Nucleotidsequenzierung verwendeten Primer wurden auf dem Oligonucleotid-synthesizer 8750 der Firma Biosearch synthetisiert (Seq. ID No. 35 + 36).

30

Primer 1: AGC AGC AGG AAG CAC TAT GG (Koordinaten aus HIV 1

Isolat HXB2: Base 7795-7814, entspricht Primer sk 68)

35

Primer 2: GAG TTT TCC AGA GCA ACC CC (Koordinaten aus HIV 1

Isolat HXB2: Base 8003-8022, entspricht Primer env b).

40 Die amplifizierte DNA wurde über ein 3 % "Nusieve"-Agarosegel (Fa. Biozyme) aufgetrennt, das amplifizierte Fragment ausgeschnitten und mit dem gleichen Volumen an Puffer (1xTBE (0,09 M TrisBorat, 0,002 M EDTA pH8,0) versetzt. Nach Inkubation des DNA-Agarosegemisches für 10 Minuten bei 70°C und nachfolgender Phenolextraktion wurde die DNA aus der wässrigen Phase durch Zugabe von 1/10 Vol 3M NaAc pH 5,5 und 2 Vol Ethanol bei -20°C 15' gefällt und anschließend in einer Zentrifuge (Eppendorf) 45 pelletiert (13000 rpm, 10', 4°C). Die pelletierte DNA wurde getrocknet, in Wasser aufgenommen und nach der photometrischen Bestimmung der DNA-Konzentration bei 260 nm im Spektralphotometer (Beckman) 50 nach der Methode von Sanger (F. Sanger, Proc. Natl. Adac. Sci., 74: 5463, 1977) sequenziert. Anstelle der Sequenzierung mit Klenow DNA Polymerase, wurde die Sequenzierungsreaktion mit einem Kit von Applied Biosystems ("Taq Dye Deoxy Terminator Cycle Sequencing", Best.-Nr.: 401150) durchgeführt. Als Primer wurden in getrennten Sequenzierungsreaktionen Primer 1 oder Primer 2 (jeweils 1 µM) eingesetzt. Die Analyse der Sequenzierungsreaktion erfolgte auf dem DNA-Sequenziergerät 373A (Applied Biosystems) nach den Vorgaben des Geräteherstellers.

Die Nucleotidsequenz des amplifizierten DNA-Bereichs und die davon abgeleitete Aminosäuresequenz ist in der Tabelle 1 dargestellt (Seq. ID No. 37-39).

Tabelle 1:

5 GCGCAGCGAACAGCGCTGACGGTACGGACCCACAGTGTACTGAAGGGTATAGTCAAC
-----+-----+-----+-----+-----+-----+
CGCGTCGCCGTTGTCGCGACTGCCATGCCTGGGTGTCACATGACTCCCATATCACGTTG
A A A T A L T V R T H S V L K G I V Q Q
10 AGCAGGACAACCTGCTGAGAGCGATAACAGGCCAGCAACACTTGCTGAGGTTATCTGTAT
-----+-----+-----+-----+-----+
TCGTCCCTGTTGGACGACTCTCGCTATGTCGGGCGTTGTGAACGACTCCAATAGACATA
15 Q D N L L R A I Q A Q Q H L L R L S V W
GGGGTATTAGACAACCTCCGAGCTCGCCTGCAAGCCTAGAAACCCCTATACAGAATCAGC
-----+-----+-----+-----+-----+
CCCCATAATCTGTTGAGGCTCGAGCGGACGTTCGGAATCTTGGGAATATGTCCTAGTCG
20 G I R Q L R A R L Q A L E T L I Q N Q Q
AACGCCTAACCTAT
-----+----- 195
25 TTGGGATTTGGATA
R L N L -

30 **Beispiel 5**

Die gefundene Nucleotidsequenz aus Tabelle 1 wurde auf homologe Sequenzen in der GENE BANK-Datenbank (Release 72, Juni 1992) mit Hilfe des GCG-Computerprogramms (Genetic Computer Group, Inc., Wisconsin USA, Version 7.1, März 1992) untersucht. In dieser Datenbank sind die meisten der bis Juli 1992 35 bekannten Nucleotidsequenzen von Immundefizienzviren humanen Ursprungs und von Isolaten aus Primaten enthalten.

Die Nucleotidsequenz aus Tabelle 1 weist im besten Fall eine 66 %ige Homologie zu einem Schimpansen-Isolat auf. Zu HIV 1-Isolaten ist MVP 5180/91 in der untersuchten DNA-Sequenz im besten Fall zu 64 % homolog. Zu HIV 2-Isolaten ist die DNA aus Tabelle 1 zu 56 % homolog. Außer der Sequenz 40 des Schimpansenisolates besteht die beste Homologie zwischen der Nucleotidsequenz aus Tabelle 1 und DNA-Abschnitten aus Primatenisolaten (SIV: Simian Immunodeficiency Virus) in einer DNA-Sequenz, die für einen Teilbereich des HÜllproteins des Isolates SIV (Afrikanische Meerkatze) TYO-1 kodiert. Die Homologie beträgt 61,5 %.

45 **Beispiel 6**

Die gefundene Aminosäuresequenz aus Tabelle 1 wurde auf homologe Sequenzen in der SWISSPROT Protein-Datenbank (Release 22, Juni 1992) mit Hilfe des GCG-Computerprogramms untersucht. In dieser Datenbank sind die meisten der bis Juni 1992 50 bekannten Proteinsequenzen von Immundefizienz-Viren humanen Ursprungs und von Isolaten aus Primaten enthalten.

Die Aminosäuresequenz aus Tabelle 1 ist im besten Fall mit 82,5 % zu einem HÜllproteinabschnitt des oben genannten Schimpansenisolates homolog. Unter HIV 1-HÜllproteinen findet man die beste Homologie mit der Aminosäuresequenz aus Tabelle 1, in dem Isolat HIV 1 Mal. Die Homologie beträgt 59 %. Zu HIV 2-HÜllproteinen beträgt die Homologie mit der Aminosäuresequenz aus Tabelle 1 im besten Fall 52 % (Isolat 55 HIV 2 Rod). Da auch HIV 1 und HIV 2-Isolate im besten Fall im korrespondierenden Proteinabschnitt nur zu 64 % identisch sind, scheint es sich bei dem Isolat MVP-5180/91 um eine HIV-Variante zu handeln, die sich deutlich von HIV 1 und von HIV 2 strukturell abgrenzt und somit einen Vertreter einer davon unabhängigen Gruppe von HIV-Viren repräsentiert.

Die Aminosäuresequenz des amplifizierten DNA-Bereiches (Tabelle 1) des HIV-Isolates MVP 5180/91 überlappt mit einem immundiagnostisch wichtigen Bereich aus dem Hüllprotein gp 41 von HIV 1 (Aminosäure 584-618*) (Tabelle 2) (Gnann et al., J. Inf. Dis. 156: 261-267, 1987; Norby et al., Nature, 329: 248-250, 1987).

5 Korrespondierende Aminosäurebereiche aus den Hüllproteinen von HIV 2 und SIV sind ebenfalls immundiagnostisch konserviert (Gnann et al., Science, S. 1348-1349, 1987). So werden Peptide aus diesem Hüllproteinbereich von HIV 1 und HIV 2 in vielen kommerziell erhältlichen HIV 1/2 Antikörper-Screeningtests als Festphasenantigene eingesetzt. Ungefähr 90 % der anti HIV 1 und anti HIV 2 positiven Seren können damit erfaßt werden.

10 Der Aminosäurebereich des MVP-5180/91-Hüllproteins (Tabelle 1) könnte aufgrund der Überlappung mit dem immundiagnostisch wichtigen Bereich aus gp 41 serodiagnostisch von Bedeutung sein. Dies wäre insbesondere dann der Fall, wenn Antiseren von HIV-infizierten Patienten mit keinem der kommerziell erhältlichen Antikörper-Screeningtests positiv reagieren würden. In diesen Fällen könnte eine Infektion mit einem MVP-5180/91 eng verwandten Virus vorliegen.

15

Tabelle 2:

20

.....RILAVERVLKDQQLLGIWGCGSGKLICTTAVPWNAS
 | : | : | . . . : | | | . :
 WGIRQLRARLQALETLIONQQRLNL.....

25 **Beispiel 7**

DNA Isolierung, Amplifizierung und strukturelle Charakterisierung von Genomabschnitten des HIV Isolates MVP-5180/91 (kodierend für gp 41)

30 Genomische DNA aus MVP-5180/91-infizierten HUT78-Zellen wurden wie beschrieben isoliert. Zur Charakterisierung von Genombereichen des Isolates MVP-5180/91 wurden PCR (Polymerase Chain Reaction)-Experimente mit Primerpaaren aus dem Hüllproteinbereich gp 41 durchgeführt. Die PCR (Saiki et al., Science 239: 487-491, 1988) und inverse PCR (Triglia et al., Nucl. Acids, Res. 16: 8186, 1988) wurden mit folgenden Modifikationen durchgeführt:

35

1. PCR

40 Für die Amplifikation HIV-spezifischer DNA Bereiche wurden 5 µl (218 µg/ml) genomicsche DNA aus MVP-5180/91 infizierten HUT78 Zellen in einem 100 µl Reaktionsansatz (0,25 mM dNTP, je 1 µM Primer 163env und Primer envend, 10 mM Tris HCl pH 8,3, 50 mM KCl, 1,5 mM MgCl₂, 0,001 % Gelatine, 2,5 units Taq Polymerase (Perkin Elmer)) pipettiert und nach folgendem Temperaturprogramm amplifiziert: 1. Initiale Denaturierung: 3 min. 95°C, 2. Amplifikation: 90 sec. 94°C, 60 sec. 56°C, 90 sec. 72°C (30 Cycles).

45

2. Inverse PCR

50 Der 5' Bereich von gp 41 (N-Terminus) und die 3' Sequenz von gp 120 wurden mittels "inverser PCR" amplifiziert. Hierzu wurden 100 µl einer genomicschen DNA Präparation (218 µg/ml) aus MVP-5180/91-infizierten HUT78-Zellen in einem Endvolumen von 200 µl mit 10 units der Restriktionsendonuklease Sau3a 1 Stunde bei 37°C verdaut. Die DNA wurde anschließend phenolisiert und mit Natriumazetat (Endkonzentration 300 mM) und 2,5 Volumen Ethanol 10 min bei -70°C gefällt, in der Eppendorfzentrifuge abzentrifugiert und das Pellet getrocknet und in 890 µl Aqua dest. resuspendiert. Nach Zugabe von 100 µl Ligasepuffer (50 mM Tris HCl, pH 7,8, 10 mM MgCl₂, 10 mM DTT, 1 mM ATP, 25 µg/ml Rinderserumalbumin) und 10 µl T4 DNA-Ligase (Fa. Boehringer, Mannheim) wurden die DNA-Fragmente 3 Stunden bei Raumtemperatur ligiert, erneut phenolisiert und mit Natriumazetat und Ethanol wie oben gefällt. Nach dem Abzentrifugieren und Trocknen wurde die DNA in 40 µl Aqua dest. resuspendiert und mit 10 units der Restriktionsendonuklease SacI (Fa. Boehringer, Mannheim) für 1 Stunde verdaut. Anschließend wurden 5 µl dieses Ansatzes in einem PCR-Experiment wie unter "1. PCR" dargestellt, eingesetzt. Anstelle der Primer

163env und envend wurden für die inverse PCR die Primer 168i und 169i verwendet.

Die Primer 163env, 168i und 169i wurden aus der bereits ermittelten Teilsequenz des HIV-Isolates MVP-5180 ausgewählt (Beispiel 4).

5 Die für die PCR/inverse PCR und die Nucleotidsequenzierung verwendeten Prim r wurd n auf dem Oligonucleotidsynthesizer 8750 der Firma Biosearch synthetisiert, wobei die Primer folgende Sequenzen aufwiesen (Sequ. ID No. 40-43):

Primer 163env: 5' CAG AAT CAG CAA CGC CTA AAC C 3'

10 Primer envend: 5' GCC CTG TCT TAT TCT TCT AGG 3'
(Position aus HIV 1 Isolat BH10:
Base 8129-8109)

Primer 168i: 5' GCC TGC AAG CCT TAG AAA CC 3'

15 Primer 169i: 5' GCA CTA TAC CCT TCA GTA CAC TG 3'

Die amplifizierte DNA wurde über ein 3 % "Nusieve"-Agarosegel (Fa. Biozyme) aufgetrennt, das amplifizierte Fragment ausgeschnitten und mit dem gleichen Volumen an Puffer (1xTBE (0.09 M-TrisBorat, 0.002 M EDTA, pH 8.0)) versetzt. Nach Inkubation des DNA-Agarosegemisches für 10 Minuten bei 70 °C und nachfolgender Phenolextraktion wurde die DNA aus der wässrigen Phase durch Zugabe von 1/10 Vol 3M NaAc, pH 5,5 und 2 Vol Ethanol bei -20 °C 15' gefällt und anschließend in einer Eppendorfzentrifuge pelletiert (13000rpm, 10', 4 °C). Die pelletierte DNA wurde getrocknet, in Wasser aufgenommen und nach der photometrischen Bestimmung der DNA-Konzentration bei 260nm im Spektralphotometer (Fa. Beckman) 20 nach der Methode von Sanger (F. Sanger, Proc. Natl. Acad. Sci., 74:5463, 1977) sequenziert. Anstelle der Sequenzierung mit Klenow DNA Polymerase wurde die Sequenzierungsreaktion mit einem Kit der Fa. Applied Biosystems ("Taq Dye Deoxy Terminator Cycle Sequencing", Best. Nr.: 401150) durchgeführt. Als Primer wurden in getrennten Sequenzierungsreaktionen Primer 163env oder Primer envend (jeweils 1µM) eingesetzt. Die amplifizierte DNA aus dem inversen PCR-Experiment wurde mit den Primern 168i und 169i 25 sequenziert. Die Analyse der Sequenzierungsreaktion erfolgte auf dem DNA-Sequenzigerät 373A (Applied Biosystems) nach den Vorgaben des Geräteherstellers.

Die Nucleotidsequenz des amplifizierten DNA-Bereichs und die davon abgeleitete Aminosäuresequenz sind in der Tabelle 3 dargestellt (Sequ. ID No. 44-46).

35

40

45

50

55

Tabelle 3

5	1	AAATGTCAAGACCAATAATAACATTACACACCCCTCACAGGGAAAAAGAGCAGTAGGAT	60
		-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+	
		TTTACAGTTCTGGTTATTATTTGTAAGTGTGGGGAGTGTCCCTTTCTCGTCATCCTA	
		M S R P I I N I H T P H R E K R A V G L	
		gp120 <-----> gp41	
10	61	TGGGAATGCTATTCTTGGGGTGCTAAGTGCAGCAGGTAGCACTATGGCGCAGCGGCAA	120
		-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+	
		ACCCCTACGATAAGAACCCCCACGATTCACGTCGTCCATCGTGATACCCGCGTCGCCGTT	
		G M L F L G V L S A A G S T M G A A A T	
15	121	CAGCGCTGACGGTACGGACCCACAGTGTACTGAAGGGTATAGTGCACAGCAGGACAACC	180
		-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+	
		GTCGCGACTGCCATGCCCTGGGTGTCACATGACTTCCCATATCACGTTGTCGTCCCTGTTGG	
		A L T V R T H S V L K G I V Q Q Q D N L	
20	181	TGCTGAGAGCGATAACAGGCCACAGCAACACTTGCTGAGGTATCTGTATGGGTATTAGAC	240
		-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+	
		ACGACTCTCGCTATGTCGGGTGTTGTGAAACGACTCCAATAGACATAACCCATAATCTG	
		L R A I Q A Q Q H L L R L S V W G I R Q	
25	241	AACTCCGAGCTGCCCTGCAAGCCTTAGAAACCCCTATACAGAACGAAACGCCTAAACC	300
		-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+	
		TTGAGGCTCGAGCGGACGTTCGGAATCTTGGGAATATGCTTAGTCGTTGCGGATTGG	
		L R A R L Q A L E T L I Q N Q Q R L N L	
30	301	TATGGGGCTGTAAGGAAAACATAATCTGTTACACATCAGTAAATGGAACACATCATGGT	360
		-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+	
		ATACCCCGACATTCCTTTGATTAGACAATGTTGAGTCACTTACCTTGTGAGTACCA	
		W G C K G K L I C Y T S V K W N T S W S	
35	361	CAGGAGGATATAATGATGACAGTATTTGGGACAACCTTACATGGCAGCAATGGGACCAAC	420
		-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+	
		GTCCTCCTATATTACTACTGTCAAAACCTGTTGGAAATGTACCGTCGTTACCTGGTTG	
		G G Y N D D S I W D N L T W Q Q W D Q H	
40	421	ACATAAAACATGTAAGCTCCATTATATATGATGAAATACAAGCAGCACAGAACAGG	480
		-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+	
		TGTATTTGTTACATCGAGGTAATATATACTACTTATGTTGTCGTTCTGGTTGTCC	
		I N N V S S I I Y D E I Q A A Q D Q Q E	
45			

50

55

EF 0391 914 A2

481 AAAAGAATGTAAGCATTGTTGGAGCTAGATGAATGGGCCCTCTCTTGGATTGGTTG 540
 TTTTCTTACATTTCTGAACACCTCGATCTACTTACCCGGAGAGAACCTTAACCAAAC
 5 K N V K A L L E L D E W A S L W N N W F D
 541 ACATAACTAAATGGTTGGTATATAAAAATAGCTATAATCATAGTGGGAGCACTAATAG 600
 TGTATTGATTACCAACACCATATTTTATCGATATTAGTATCACCCTCGTATTAC
 10 I T K W L W Y I K I A I I I V G A L I G
 601 GTATAAGAGTTATCATGATAGTACTTAATCTAGTGAAGAACATTAGGCAGGGATATCAAC 660
 CATATTCTCAATAGTACTATCATGAATTAGATCACTTCTGTAATCCGTCCTATAGTTG
 15 I R V I M I V L N L V K N I R Q G Y Q P
 561 CCCTCTCGTTGCAGATCCCTGTCCACACCGGAGGAAGCAGAACGCCAGGAAGAACAG 720
 20 GGGAGAGCAACGTCTAGGGACAGGGTGTGGCCCTTCGTCTTGCCTTCGGTCTTGTGTC
 L S L Q I P V P H R Q E A E T P G R T G
 721 GAGAAGAAGGTGGAGAAGGAGACAGGCCAAGTGGACAGCCTTGCACCAGGATTCTGC 780
 CTCTTCTTCCACCTCTCCTGTCCGGTTACCTGTGCGAACGGTGGTCTAAGAACG
 25 E E G G E G D R P K W T A L P P G F L Q
 781 AACAGTTGTACACGGATCTCAGGACAATAATCTTGTGGACTTACCACTCTTGAGCAACT 840
 TTGTCAACATGTGCCTAGACTCTGTATTAGAACACCTGAATGGTGGAGAACCTGTTGA
 30 Q L Y T D L R T I I L W T Y H L L S N L
 841 TAATATCAGGGATCCGGAGGCTGATCGACTACCTGGACTGGACTGTGGATCCTGGAC 900
 35 ATTATAGTCCCTAGGCCTCCGACTAGCTGATGGACCCCTGACACCTAGGACCCCTG
 I S G I R R L I D Y L G L G L W I L G Q
 901 AAAAGACAATTGAAGCTTGTAGACTTGTGGAGCTGTAATGCAATATTGGCTACAAGAAT 960
 TTTTCTGTTAACCTCGAACATCTGAAACACCTCGACATTACGTTATAACCGATGTTCTTA
 40 K T I E A C R L C G A V M Q Y W L Q E L
 961 TGAAAAATAGTGTACAAACCTGCTTGTGATACTATTGCACTGTCAGTTGCCAATTGGACTG 1020
 ACTTTTATCACGATGTTGGACGAACTATGATAACGTCAAGTCAGTCAACGGTTAACCTGAC
 45 K N S A T N L L D T I A V S V A N W T D
 50
 1021 ACGGCATCATCTTAGGTCTACAAAGAATAGGACAAGG 1057
 TGCCGTAGTAGAATCCAGATGTTCTTACCTGTTCC
 55 G I I L G L Q R I G Q

B Beispiel 8

Die gefundene Nucleotidsequenz aus Tabelle 3 wurde auf homologe Sequenzen in der GENE BANK-Datenbank (Release 72, Juni 1992) mit Hilfe des GCG-Computerprogramms (Genetic Computer Group, Inc. 5 Wisconsin USA, Version 7.1, März 1992) untersucht. In dieser Datenbank sind die meisten der bis Juli 1992 bekannten Nucleotidsequenzen von Immundefizienzviren humanen Ursprungs und von Isolaten aus Primaten enthalten.

Die Nucleotidsequenz aus Tabelle 3 weist im besten Fall eine 62 %ige Homologie zu einem HIV 1-Isolat auf. Zu HIV 2 Isolaten ist die DNA aus Tabelle 5 zu 50 % homolog.

10 Die aus der Nucleotidsequenz aus Tabelle 3 abgeleitete Aminosäuresequenz wurde auf homologe Sequenzen in der SWISSPROT Protein-Datenbank (Release 22, Juni 1992) mit Hilfe des GCG-Computerprogramms untersucht. In dieser Datenbank sind die meisten der bis Juni 1992 bekannten Proteinsequenzen von Immundefizienzviren humanen Ursprungs und von Isolaten aus Primaten enthalten.

15 Die Aminosäuresequenz aus Tabelle 3 ist im besten Fall mit 54 % zu dem korrespondierenden HÜllproteinabschnitt eines Schimpansen-Isolates CIV (SIVcpz) homolog und zu 54,5 % zu dem HIV 1-Isolat Mai. Zu HIV 2-HÜllproteinen beträgt die Homologie mit der Aminosäuresequenz aus Tabelle 3 im besten Fall 34 % (Isolat HIV 2 D194).

20 Vergleicht man demgegenüber die gp 41-Aminosäuresequenz von HIV 1 mit den in der SWISSPROT-Datenbank vorhandenen HIV 1 gp 41-Sequenz, ergeben sich wie erwartet im besten Fall eine fast 100 %ige Homologie und im schlechtesten Fall eine 78 %ige Homologie.

25 Aufgrund dieser deutlichen strukturellen Unterschiede zwischen dem Sequenzbereich aus Tabelle 3 und dem korrespondierenden Abschnitt aus HIV 1 und HIV 2 scheint es sich bei dem Isolat MVP-5180/91 um eine HIV-Variante zu handeln, die sich von HIV 1 und HIV 2 deutlich strukturell abgrenzt. Möglicherweise ist MVP-5180/91 einer eigenen, von HIV 1 und HIV 2 sich abgrenzenden Gruppe von HIV-Viren zuzuordnen.

26 Das Peptid von Aminosäure 584-618 des HIV 1-HÜllproteinbereichs ist von besonderem serodiagnostischem Interesse (Numerierung nach Wain Hobson et al., Cell 40:9-17, 1985; Gnann et al., J. Inf. Dis. 156:261-267, 1987; Norby et al., Nature, 329:248-250, 1987). Korrespondierende Aminosäurebereiche aus den HÜllproteinen von HIV 2 und SIV sind ebenfalls immundiagnostisch konserviert (Gnann et al., Science, S. 1346-1349, 1987). So werden Peptide aus diesem HÜllproteinbereich von HIV 1 und HIV 2 in vielen 30 kommerziell erhältlichen HIV 1/2 Antikörper-Screeningtests als Festphasenantigene eingesetzt. Ungefähr 99 % der Anti-HIV 1 und Anti-HIV 2 positiven Seren können damit erfaßt werden.

35 Der korrespondierende Aminosäurebereich des MVP-5180/91-HÜllproteins (Tabelle 4) als auch das gesamte gp 41 dieses Isolates könnten serodiagnostisch insbesondere dann von Bedeutung sein, wenn Antiseren von HIV-infizierten Patienten in kommerziell erhältlichen Antikörper-Screening-Tests nur schwach oder überhaupt nicht reagieren würden. In diesen Fällen könnte eine Infektion mit einem MVP-5180/91 eng verwandten Virus vorliegen.

Tab. 4:

40

1 RILAVERYLKQQQLLGIGCNSGKLICTAVPWNAS

2 LQ L TLIQN R NL K Y S K T

45

1 HIV 1 Aminosäuresequenz aus gp41

2 MvP5180-Sequenz aus gp41. Nur Unterschiede zu der HIV 1-Sequenz sind ausgedruckt.

50

Das Peptid, das mit Hilfe der von MvP 5180 stammenden Information aufgefunden wurde, hat also die Aminosäuresequenz: RLOALETLIQNQQRLNLWGCKGKLYTSVKWNTS.

55 Gegenstand der vorliegenden Erfindung sind daher Peptide, die rekombinant oder synthetisch hergestellt werden können und die oben angegebene Sequenz oder Teilsequenz aufweisen, wobei die Teilsequenzen wenigstens 6 aufeinanderfolgende Aminosäuren, bevorzugt 9 und besonders bevorzugt 12 aufeinanderfolgende Aminosäuren aufweisen.

B Beispiel 9

Klonierung des Gesamtgenoms des HIV-Isolates MvP5180

5 a) Herstellung einer genomischen Bibliothek

Genomische DNA aus MvP5180-infizierten HUT78-Zellen wurde wie beschrieben isoliert.

300 µg dieser DNA wurde in einem Volumen von 770 µl mit 0,24 U des Restriktionsenzymes Sau3A für 45 min inkubiert. Die dadurch nur partiell geschnittene DNA wurde anschließend über ein 0,7 %iges Agarose-
10 gel (low melting agarose, Nusieve) größenzfraktioniert und Fragmente zwischen 10 und 21 kb ausgeschnitten. Die Agarose wurde für 10 min bei 70 °C geschmolzen und mit demselben Volumen Puffer (1 x TBE, 0,2 M NaCl) versetzt. Anschließend folgte nach zweimaliger Phenol- und einmaliger Chloroformextraktion die Fällung der DNA durch Zugabe von 1/10 Vol 3 M Natriumacetatlösung (pH 5,9) und 2,5 Vol Ethanol bei
15 -70 °C für 10 min. Die gefällte DNA wurde abzentrifugiert, getrocknet und in einer Konzentration von 1 µg/ml in Wasser gelöst.

Die Ausbeute an größenzfraktionierter DNA betrug etwa 60 µg. 5 µg dieser DNA wurden mit 1 U Alkalische Phosphatase im entsprechenden Puffer für 20 min bei 37 °C inkubiert. Durch Abspaltung des 5'-terminalen Phosphatrestes wurden so multiple Insertionen von größenzfraktionierter DNA vermindert. Die Phosphatase-
20 Behandlung wurde durch Phenolisierung gestoppt, die DNA wie oben gefällt und zusammen mit 1 µg des Vektors (2 DASH, BamHI geschnitten, Stratagene Nr.: 247611) in 6 µl Gesamtvolume mit 2 Weiss-Units Lambda T4 Ligase für 12 Stunden bei 15 °C ligiert. Nach erfolgter Ligation wurde die DNA mit Hilfe eines Verpackungskits (Gigapack II Gold, Stratagene Nr.: 247611) genau nach Angaben des Herstellers in Phagenhüllen verpackt.

25 b) Radioaktive Markierung der DNA-Probe

Für die Markierung wurde der "Random Primed DNA Labeling Kit" von Boehringer Mannheim (Nr.: 713 023) eingesetzt. Markiert wurde das PCR-Produkt, welches wie in Beispiel 3 beschrieben mit den Primern sk68 und envb erhalten wurde. 1 µg dieser DNA wurde durch 2 x 5 min Kochen und anschließender Abkühlung in Eiswasser denaturiert. Zur Markierung wurden 50 mCi [α -³²P]-dCTP (NEN, Nr.: NEX-053H) zugegeben. Sonstige Zusätze wurden nach Herstellerangaben pipettiert. Nach einer Inkubation von 30 min bei 37 °C erfolgte eine Fällung der nun radioaktiv markierten DNA.

35 c) Screening der Phagen-Bibliothek

Zu 200 µl einer bei 30 °C über Nacht angezogenen Kultur (Stamm SRB(P2) [Stratagene, Nr.: 247611] in LB-Medium, welches 10 mM MgSO₄, sowie 0,2 % Maltose enthielt) wurden 20000 pfu (Plaque forming units) der Bibliothek in 100 µl SM-Puffer (5,8 g NaCl, 2 g MgSO₄, 50 ml 1 M Tris, pH 7,5 und 5 ml einer 2 % Gelatinelösung in 1 l H₂O gekösst) gegeben, die Phagen 20 min bei 37 °C an die Bakterien adsorbiert, mit
40 7,5 ml auf 55 °C abgekühlter Top-Agarose gemischt und auf einer vorgewärmmten Lb-Agarplatte von 14 cm Durchmesser verteilt. Nach etwa 8 Stunden erreichten die Plaques Konfluenz. Daraufhin wurden Nitrocellulosefilter für wenige Minuten auf die Platten gelegt und asymmetrische Markierungen angebracht. Nach vorsichtigem Abheben wurden die Filter für 2 min denaturiert (0,5 M NaOH, 1,5 M NaCl) und dann 5 min neutralisiert (0,5 M Tris, pH 8, 1,5 M NaCl). Nach anschließendem Backen der Filter bei 80 °C für 60 min,
45 konnten die Filter mit der Probe hybridisiert werden.

Zur Vorhybridisierung wurden die Filter in 15 ml Hybridisierungslösung (50 % Formamid, 0,5 % SDS, 5 x SSPE, 5 x Denhardt's Lösung und 0,1 mg/ml Lachssperma-DNA) pro Filter bei 42 °C unter Schütteln 2-3 h inkubiert. Die [³²P]-markierten DNA-Proben wurden 2-5 min bei 100 °C denaturiert, auf Eis abgekühlt, der Vorhybridisierungslösung zugesetzt und 12 Stunden bei 42 °C hybridisiert. Anschließend wurden die Filter
50 bei 60 °C zunächst mit 2 x SSC/0,1 % SDS, dann mit 0,2 x SSC/0,1 % SDS gewaschen. Nach Trocknen der Filter wurden Hybridisierungssignale mit Hilfe des Röntgenfilms X-OMAT™ AR (Kodak) detektiert. Die Plaques, denen ein Signal zugeordnet werden konnte, wurden nach Elution in SM-Puffer in weiteren Verdünnungsschritten vereinzelt.

Der unten beschriebene Klon konnte nach dem Screening von 2 x 10⁶ Plaques identifiziert werden.

d) Isolierung der Phagen-DNA und Subklonierung

Mit 10 μ l eines Phageneluats in SM-Puffer, wurde eine Übernachtkultur des Wirtsstammes SRB (P2) so infiziert, daß nach zunächst dichtem Wachstum der Kultur nach etwa 6-8 h Lyse erfolgte. Von der lysierten Kultur wurden Zellreste durch zweimalige Zentrifugation bei 8000 g für 10 min abgetrennt. Anschließend wurden die Phagen durch Zentrifugation pelletiert (35000 g, 1 h), in 700 μ l 10 mM MgSO₄ aufgenommen und solange phenolisiert, bis keine Proteininterphase mehr zu sehen war. Daraufhin wurde die Phagen-DNA gefällt, mit dem Restriktionsenzym EcoRI geschnitten und die daraus erhaltenen EcoRI-Fragmente in den Vektor Bluescript KS⁻ (Stratagene, Nr.: 212208) subkloniert. Insgesamt wurden 4 Klone erhalten:

10

Plasmid	Beginn ¹	Ende ¹
pSP1	1	1785
pSP2	1786	5833
pSP3	5834	7415
pSP4	7680	9793

15

¹bzgl. der folgenden Gesamtsequenz

20 Das fehlende Stück zwischen Base 7416 und 7659 wurde durch PCR mit den Primern 157 (CCA TAA TAT TCA GCA GAA CTA G) und 226 (GCT GAT TCT GTA TAA GGG) erhalten. Als DNA-Template wurde die Phagen-DNA des Klons verwendet. Die Bedingungen für die PCR waren: 1.) Initiale Denaturierung: 94 °C, 3 min, 2.) Amplifikation: 1.5 min 94 °C, 1 min 56 °C und 1 min 72 °C für 30 Zyklen. Die Sequenzierung der DNA erfolgte wie in Beispiel 4 beschrieben. Vom gesamten Genom wurde sowohl 25 der Strang- als auch der Gegenstrang sequenziert. Bei allen EcoRI Schnittstellen wurde durch PCR mit Phagen-DNA des Klons als DNA-Template verifiziert, daß es sich an den Subklonübergängen jeweils um singuläre EcoRI-Schnittstellen handelt.

Tab. 5

30

Die Lage der Gene der Virusproteine GAG, POL und ENV in der Gesamtsequenz von MvP5180		
Gen	Start ¹	Stop ¹
GAG	817	2310
POL	2073	5153
ENV	6260	8887

35 1.) Die Zahlen geben die Basenpositionen in der Gesamtsequenz von MvP5180/91

40

Die Gesamtsequenz von MvP5180/91 ist in Fig. 4 dargestellt.

Beispiel 10

45

Abgrenzung der Gesamtsequenz von MvP5180/91 von anderen HIV1-Isolaten

Grundlage für die folgenden Sequenzvergleiche war die Datenbanken Genbank Release 75 von 2.93, EMBL 33 von 12.92 und Swissprot 24 von 1.93. Homologievergleiche erfolgten mit der GCG-Software (Version 7.2, 10.92. der Genetics-Computer-Group, Wisconsin).

50

Zunächst wurden auf Aminosäureebene die Sequenzen von GAG, POL und ENV mit dem Programm "Wordsearch" mit der Datenbank verglichen. Die 50 besten Homologen wurden mit dem Programm "Pileup" jeweils untereinander verglichen. Daraus geht deutlich hervor, daß MvP5180/91 in den HIV1-Stammbaum fällt, dort aber sehr früh, sogar noch vor dem Schimpansenvirus SIVcpz abzweigt, also eine neue Subfamilie von HIV1 repräsentiert. Um Zahlenwerte für die Homologien zu erhalten, wurde mit dem Programm "Gap" MvP5180 mit den jeweils am besten passenden HIV1, HIV2 und SIV-Sequenzen und zusätzlich mit der SIVcpz-Sequenz verglichen.

Tab. 6

Homologiewerte der Aminosäuresequenz von GAG, POL und ENV des MvP5180/91-Is. lat. s.								
	GAG	SIVcpz	70,2% 83,6%	HIV1u ²	69,9% 81,2%	HIV2d ³	53,6% 71,3%	SIV1a ⁴
	POL	SIVcpz	78,0% 88,0%	HIV1u ²	76,1% 86,8%	HIV2d ³	57,2% 71,9%	SIVgb ⁵
	ENV	SIVcpz	53,4% 67,1%	HIV1h ¹	50,9% 67,2%	HIV2d ³	34,4% 58,7%	SIVat ⁶
5								
10								

¹h = hz321/Zaire,²u = u455/Uganda,³d = jrcst,⁴a = agm155,⁵gb = gb1⁶at = agm

20 Der obere Zahlenwert drückt die Identität, der untere die Ähnlichkeit beider Sequenzen aus.
 Weiter wurde die Datenbank mit "Wordsearch" und "Gap" auf Nukleotidebene durchsucht. Die Homologiewerte für die jeweils besten "matches" sind in Tabelle 7 zusammengefaßt.

Tab. 7

Homologiewerte der Nukleotidsequenz von MvP5180/91					
	HIV1		HIV2		
25	gag	HIVelicg	70,24 %	HIV2bihz	60,0 %
30	pol	HIVmal	75,0 %	HIV2cam2	62,9 %
35	env	HIVsimi84	59,7 %	HIV2gha	49,8 %

Beispiel 11

Beschreibung der PCR Amplifizierung, Klonierung und Sequenzierung des gag Gens des HIV 5180 Isolates

40 Um die im Laufe der Virusvermehrung auftretenden Spontanmutationen darzustellen, wurde ein Teil des Virusgenoms mit der PCR-Technik kloniert und die so erhaltene DNS-Sequenz verglichen mit der Sequenz gemäß Fig. 4.

45 Die gag Sequenz wurde vom LTR ("long terminal repeat", LTR1 primer) des linken Endes des MvP 5180 Genomes bis in das pol (Polymerase Gen, pol3.5i primer) hinein überlappend kloniert. Die Klonierungsstrategie ist in Fig. 5 schematisch dargestellt.

46 Die PCR Reaktionen wurden mit den u.g. DNA Primern, deren Sequenzen von der HIV-1 Konsensus-Sequenz abgeleitet wurden, durchgeführt. Die Sequenzierungen erfolgten mit Hilfe der Dideoxykettenabbruchmethode.

50 Die für das MvP 5180 gag Gen kodierende Sequenz erstreckt sich von Nukleotid 817 (A des ATG Startkodons) bis Nukleotid 2300 (A des letzten Kodons) (Sequ. ID No. 47-53).

5 LTR1: 5' - CTA GCA GTG GCG CCC GAA CAG G -3'
 gag3.5: 5' - AAT GAG GAA GCU GCA GAU TGG GA -3' (U=A/T)
 gag 3.5i: 5' - TCC CAU TCT GCU GCT TCC TCA TT -3' (U=A/T)
 gag5: 5' - CCA AGG GGA AGT GAC ATA GCA GGA AC -3'
 gag959: 5' - CGT TGT TCA GAA TTC AAA CCC -3'
 gag11i: 5' - TCC CTA AAA AAT TAG CCT GTC -3'
 10 pol3.5i: 5' - AAA CCT CCA ATT CCC CCT A -3'
 15

10 Die bei der PCR-Technik erhaltene DNS-Sequenz wurde der in Figur 4 dargestellten DNS-Sequenz gegenübergestellt. Ein Vergleich der beiden DNS-Sequenzen ist in Figur 8 dargestellt. Hierbei wurde festgestellt, daß sich die Nukleotide in ca. 2 % voneinander unterscheiden, obwohl es sich um dasselbe Virus handelt. In Fig. 8 stellt jeweils die obere Zeile die DNS-Sequenz dar, die in Fig. 4 dargestellt ist und die untere Zeile stellt die mit PCR-Technik erhaltene DNS-Sequenz dar.

15 Weiterhin wurde die Aminosäuresequenz des mit PCR-Technik ermittelten Proteins gag der Aminosäuresequenz des aus Fig. 4 abgeleiteten entsprechenden Proteins gegenübergestellt. Dabei wurde ein Unterschied der Aminosäure von ca. 2,2 % ermittelt. Der Vergleich ist in Fig. 7 dargestellt, wobei die untere Zeile jeweils die Aminosäuresequenz darstellt, die von der mit PCR-Technik erhaltenen Sequenz abgeleitet wurde.

20

Beispiel 12

Es wurde die Sequenz des erfindungsgemäßen Virus MvP 5180 verglichen mit den Konsensus-Sequenzen von HIV1 und HIV2 und soweit bekannt mit der Sequenz von ANT-70 (WO 89/12094).

25 Dabei wurden folgende Ergebnisse erhalten:

Tab. 8

Genort	abweichende Nukleotide	Zahl der Nukleotide	% Homologie (genähert)
LTR	207	630	HIV-1 67 %
	308		HIV-2 51 %
	115		ANT 70 82 %
GAG	448	1501	HIV-1 70 %
	570		HIV-2 62 %
POL	763	3010	HIV-1 74 %
	1011		HIV-2 66 %
VIF	183	578	HIV-1 68 %
	338		HIV-2 42 %
ENV	1198	2534	HIV-1 53 %
	1289		HIV-2 49 %
NEF	285	621	HIV-1 54 %
	342		HIV-2 45 %
total	3082	8874	HIV-1 65 %
	3858		HIV-2 58 %

50 In der obigen Tabelle bedeuten "HIV-1" Konsensus-Sequenzen von HIV-1 Viren; "HIV-2" Konsensus-Sequenzen von HIV-2 Viren; ANT-70 aus der WO 89/12094 bekannte Teilsequenz eines als HIV-3 bezeichneten Virus.

55 Gegenstand der vorliegenden Erfindung sind daher Viren, DNS-Sequenzen, Aminosäuresequenzen sowie Teilsequenzen davon, die eine solche Homologie mit der in Fig. 4 dargestellten Sequenz aufweisen, bezogen auf die Genorte, daß höchstens die in Tabelle 9 angegebenen Anteile, ausgedrückt in %-Werten, unterschiedlich sind.

Tab. 9

H mologl. b. zog. n auf G norte, ausgedrückt als maximale Unterschied				
	Genort	Unterschiede	bevorzugte Unterschiede	besonders bevorzugte Unterschiede
5	LTR	17 %	15 %	10 %
	GAG	29 %	28 %	14 %
	POL	25 %	24 %	12 %
	VIF	31 %	30 %	15 %
	ENV	46 %	45 %	22 %
	NEF	16 %	12 %	10 %

15 Die in Tabelle 9 angegebenen Homologiewerte in % bedeuten, daß bei einer Gegenüberstellung der Sequenz gemäß Fig. 4 mit einer Sequenz eines anderen Virus höchstens ein den oben angegebenen Prozentwerten entsprechender Anteil der Sequenz unterschiedlich sein darf.

Beispiel 13

V3-loop / V3-Schlaufe

20 Diese Schlaufe ist die hauptneutralisierende Region im HIV und die Dokumentation der immunologischen Spezifitäten der Region sind in der Figur 8 zusammengefaßt. Dies ist eine Kopie aus AIDS aus einer Arbeit von Peter Nara (1990).

25 Die V3-Schlaufe ist dann in Aminosäuren-Ebene aufgezeichnet und mit dem IIIB Virus jetzt LAI und dem ersten HIV-2 Isolat (ROD) verglichen. Einzelne Aminosäuren an der Cystin-Brücke sind konserviert. Während die Krone von HIV-1 GPGR oder GPGQ ist und die von HIV-2 GHVF ist die Krone des MvP5180/91 gebildet aus den Aminosäuren GPMR. Das Motiv mit dem Methionin ist bisher nicht 30 beschrieben worden und unterstreicht die Individualität des MvP 5180/91.

Nachdem die Nukleinsäuresequenz des Virus ermittelt war, wurde der V3-Loop-Bereich mit Hilfe der PCR-Technik unter Verwendung geeigneter Primer amplifiziert. Hierbei konnten Mutationen beobachtet werden, insbesondere eine Veränderung des Methionin-Kodons (ATG) zu einem Leucin-Kodon (CTG).

35 Nachfolgend wird eine Gegenüberstellung der von der klonierten Nukleinsäure abgeleiteten Aminosäuresequenz und der Sequenz gegeben, die nach Amplifizierung mittels PCR-Technologie erhalten wurde (Sequ. ID No. 54-55):

MvP 5180 (kloniert):

40 CIREGIAEVQDIYTGPMRWRSMTLKRSNNNTSPRSRVAYC

MvP 5180 (PCR-Technik):

45 CIREGIAEVQDLHTGPLRWRSMTLKKSSNSHTQPRSKVAYC

Beispiel 14

50 Um zu zeigen, daß mit Hilfe des erfindungsgemäßen Virus MvP 5180 bzw. davon abgeleiteten Antigenen auch solche Seren als HIV-1 positiv nachgewiesen werden können, die bei Verwendung eines normalen HIV-1+2 Screening-Tests nicht erfaßt werden können, wurden verschiedene Seren von Patienten aus Kamerun im EIA-Test überprüft.

55 Bei einer Studie in Kamerun wurden 158 Anti-HIV-1-positive Seren überprüft. Bei zwei dieser Seren wurden erhebliche, diagnostisch relevante Unterschiede beobachtet. Bei der nachfolgenden Tabelle 10 werden die gemessenen Extinktionen angegeben. CAM-A bzw. CAM-B stehen für die Seren verschiedener Patienten.

Tabelle 10

Patientenserien	MvP 5180-EIA	HIV-1 + HIV-2 EIA
CAM-A	2.886	1.623
CAM-B	1.102	0.386

Der Cutoff für beide Tests betrug 0.300.

In einer weiteren Studie mit 47 Anti-HIV-1-positiven Seren aus Kamerun fielen zwei Seren besonders auf. Eines davon (93-1000) stammt von einem wenig symptomatischen, das andere (93-1001) von einem AIDS-kranken Patienten. In der folgenden Tabelle 11 werden die Extinktionswerte der beiden EIA-Tests einander gegenübergestellt:

Tabelle 11

Patientenserum	MvP 5180-EIA	HIV-1 + HIV-2 EIA
93-1000	> 2.5	1.495
93-1001	0.692	0.314

Auch hier betrug der Cutoff 0.3. Die Extinktionswerte des Patienten 93-1001 zeigen, daß der normale HIV-1 + HIV-2 EIA versagen kann, wohingegen durch Einsatz des erfindungsgemäßen Antigens ein klarer Nachweis möglich ist.

SEQUENZPROTOKOLL

(1) ALLGEMEINE INFORMATION:

ANMELDER:

- (A) NAME: Behringwerke Aktiengesellschaft
- (B) STRASSE: Postfach 11 40
- (C) ORT: Marburg
- (E) LAND: Germany
- (F) POSTLEITZAHL: 35001

ANMELDETITEL:

Retrovirus aus der HIV-Gruppe und
dessen Verwendung

ANZAHL DER SEQUENZEN: 60

COMPUTER-LESBARE FORM:

- (A) DATENTRÄGER: Floppy Disk
- (B) COMPUTER: IBM PC compatible
- (C) BETRIEBSSYSTEM: PC-DOS/MS-DOS
- (D) SOFTWARE: PADAT Sequenzmodul Version 1.0

50

55

(2) INFORMATION ZU SEQ ID NO: 1:

5 (i) SEQUENZ CHARAKTERISTIKA:
 (A) LÄNGE: 18 Basenpaare
 (B) ART: Nukleinsäure
 (C) STRANGFORM: Einzel
 (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: synthetische DNA

10 (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 1:

CTACTAGTAC CCTTCAGG

18

15

(2) INFORMATION ZU SEQ ID NO: 2:

20 (i) SEQUENZ CHARAKTERISTIKA:
 (A) LÄNGE: 21 Basenpaare
 (B) ART: Nukleinsäure
 (C) STRANGFORM: Einzel
 (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: synthetische DNA

25

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 2:

CGG TCT ACA TAG TCT CTA AAG

21

30

(2) INFORMATION ZU SEQ ID NO: 3:

35 (i) SEQUENZ CHARAKTERISTIKA:
 (A) LÄNGE: 21 Basenpaare
 (B) ART: Nukleinsäure
 (C) STRANGFORM: Einzel
 (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: synthetische DNA

40

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 3:

CCACCTATCC CAGTAGGAGA A

21

45

50

55

22

(2) INFORMATION ZU SEQ ID NO: 4:

5 (i) SEQUENZ CHARAKTERISTIKA:
 (A) LÄNGE: 30 Basenpaare
 (B) ART: Nukleinsäure
 (C) STRANGFORM: Einzel
 (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: synthetische DNA

10 (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 4:

CCTTTGGTCC TTGTCTTATG TCCAGAACG

30

15

(2) INFORMATION ZU SEQ ID NO: 5:

20 (i) SEQUENZ CHARAKTERISTIKA:
 (A) LÄNGE: 25 Basenpaare
 (B) ART: Nukleinsäure
 (C) STRANGFORM: Einzel
 (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: synthetische DNA

25

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 5:

TGGGAAGTTC AATTAGGAAT ACCAC

25

30

(2) INFORMATION ZU SEQ ID NO: 6:

35 (i) SEQUENZ CHARAKTERISTIKA:
 (A) LÄNGE: 26 Basenpaare
 (B) ART: Nukleinsäure
 (C) STRANGFORM: Einzel
 (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: synthetische DNA

40

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 6:

CCTACATAGA AATCATCCAT GTATTG

26

45

50

55

(2) INFORMATION ZU SEQ ID NO: 7:

5 (i) SEQUENZ CHARAKTERISTIKA:
 (A) LÄNGE: 19 Basenpaare
 (B) ART: Nukleinsäure
 (C) STRANGFORM: Einz 1
 (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: synthetische DNA

10 (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 7:

TGGATGTGGG TGATGCATA

19

15

(2) INFORMATION ZU SEQ ID NO: 8:

20 (i) SEQUENZ CHARAKTERISTIKA:
 (A) LÄNGE: 21 Basenpaare
 (B) ART: Nukleinsäure
 (C) STRANGFORM: Einzel
 (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: synthetische DNA

25 (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 8:

AGCACATTGT ACTGATATCT A

21

30

(2) INFORMATION ZU SEQ ID NO: 9:

35 (i) SEQUENZ CHARAKTERISTIKA:
 (A) LÄNGE: 22 Basenpaare
 (B) ART: Nukleinsäure
 (C) STRANGFORM: Einzel
 (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: synthetische DNA

40 (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 9:

AGTGGGGGGA CATCAAGCAG CC

22

45

(2) INFORMATION ZU SEQ ID NO: 10.

50 (i) SEQUENZ CHARAKTERISTIKA:
 (A) LÄNGE: 22 Basenpaare
 (B) ART: Nukleinsäure
 (C) STRANGFORM: Einzel
 (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: synthetische DNA

55 (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 10:

TGCTATGTCA CTTCCCCCTTG GT

22

(2) INFORMATION ZU SEQ ID NO: 11:

5 (i) SEQUENZ CHARAKTERISTIKA:
 (A) LÄNGE: 22 Basenpaare
 (B) ART: Nukleinsäure
 (C) STRANGFORM: Einzel
 (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: synthetische DNA

10 (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 11:

CCATGCAAAT GTTAAAGAG AC

22

15

(2) INFORMATION ZU SEQ ID NO: 12:

20 (i) SEQUENZ CHARAKTERISTIKA:
 (A) LÄNGE: 19 Basenpaare
 (B) ART: Nukleinsäure
 (C) STRANGFORM: Einzel
 (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: synthetische DNA

25

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 12:

GGCCTGGTGC AATAGGCC

19

30

(2) INFORMATION ZU SEQ ID NO: 13:

35 (i) SEQUENZ CHARAKTERISTIKA:
 (A) LÄNGE: 20 Basenpaare
 (B) ART: Nukleinsäure
 (C) STRANGFORM: Einzel
 (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: synthetische DNA

40

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 13:

GTGCTTCCAC AGGGATGGAA

20

45

(2) INFORMATION ZU SEQ ID NO: 14:

50 (i) SEQUENZ CHARAKTERISTIKA:
 (A) LÄNGE: 18 Basenpaare
 (B) ART: Nukleinsäure
 (C) STRANGFORM: Einzel
 (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: synthetische DNA

55

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 14:

ATCATCCATG TATTGATA

18

(2) INFORMATION ZU SEQ ID NO: 15:

5 (i) SEQUENZ CHARAKTERISTIKA:
 (A) LÄNGE: 20 Basenpaare
 (B) ART: Nukleinsäure
 (C) STRANGFORM: Einzel
 (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: synthetische DNA

10

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 15:

AATGGAGCCA GTAGATCCTA

20

15

(2) INFORMATION ZU SEQ ID NO: 16:

20 (i) SEQUENZ CHARAKTERISTIKA:
 (A) LÄNGE: 20 Basenpaare
 (B) ART: Nukleinsäure
 (C) STRANGFORM: Einzel
 (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: synthetische DNA

25

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 16:

TGTCTCCGCT TCTTCCTGCC

20

30

(2) INFORMATION ZU SEQ ID NO: 17:

35 (i) SEQUENZ CHARAKTERISTIKA:
 (A) LÄNGE: 20 Basenpaare
 (B) ART: Nukleinsäure
 (C) STRANGFORM: Einzel
 (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: synthetische DNA

40

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 17:

GAGCCCTGGA AGCATCCAGG

20

45

(2) INFORMATION ZU SEQ ID NO: 18:

50 (i) SEQUENZ CHARAKTERISTIKA:
 (A) LÄNGE: 20 Basenpaare
 (B) ART: Nukleinsäure
 (C) STRANGFORM: Einzel
 (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: synthetische DNA

55

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 18:

GGAGATGCCT AAGGCTTTG

20

(2) INFORMATION ZU SEQ ID NO: 19:

5 (i) SEQUENZ CHARAKTERISTIKA:
 (A) LÄNGE: 17 Basenpaare
 (B) ART: Nukleinsäure
 (C) STRANGFORM: Einzel
 (D) TOPOLOGIE: linear

10 (ii) ART DES MOLEKÜLS: synthetische DNA

15 (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 19:

TGTTCCCTTGG GTTCTTG

17

(2) INFORMATION ZU SEQ ID NO: 20:

20 (i) SEQUENZ CHARAKTERISTIKA:
 (A) LÄNGE: 20 Basenpaare
 (B) ART: Nukleinsäure
 (C) STRANGFORM: Einzel
 (D) TOPOLOGIE: linear

25 (ii) ART DES MOLEKÜLS: synthetische DNA

26 (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 20:

GAGTTTTCCA GAGCAACCCC

20

(2) INFORMATION ZU SEQ ID NO: 21:

35 (i) SEQUENZ CHARAKTERISTIKA:
 (A) LÄNGE: 20 Basenpaare
 (B) ART: Nukleinsäure
 (C) STRANGFORM: Einzel
 (D) TOPOLOGIE: linear

40 (ii) ART DES MOLEKÜLS: synthetische DNA

45 (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 21:

AGCAGCAGGA AGCACTATGG

20

(2) INFORMATION ZU SEQ ID NO: 22:

50 (i) SEQUENZ CHARAKTERISTIKA:
 (A) LÄNGE: 24 Basenpaare
 (B) ART: Nukleinsäure
 (C) STRANGFORM: Einzel
 (D) TOPOLOGIE: linear

55 (ii) ART DES MOLEKÜLS: synthetische DNA

55 (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 22:

GCCCCAGACT GTGAGTTGCA ACAG

24

(2) INFORMATION ZU SEQ ID NO: 23:

5 (i) SEQUENZ CHARAKTERISTIKA:
 (A) LÄNGE: 22 Basenpaare
 (B) ART: Nukleinsäure
 (C) STRANGFORM: Einzel
 (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: synthetische DNA

10 (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 23:

GCACAGTACA ATGTACACAT GG

22

15

(2) INFORMATION ZU SEQ ID NO: 24:

20 (i) SEQUENZ CHARAKTERISTIKA:
 (A) LÄNGE: 22 Basenpaare
 (B) ART: Nukleinsäure
 (C) STRANGFORM: Einzel
 (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: synthetische DNA

25

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 24:

CAGTAGAAAA ATTCCCTTCC AC

22

30

(2) INFORMATION ZU SEQ ID NO: 25:

35 (i) SEQUENZ CHARAKTERISTIKA:
 (A) LÄNGE: 31 Basenpaare
 (B) ART: Nukleinsäure
 (C) STRANGFORM: Einzel
 (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: synthetische DNA

40

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 25:

TCAGGATCCA TGGGCAGTCT AGCAGAAGAA G

31

45

(2) INFORMATION ZU SEQ ID NO: 26:

50 (i) SEQUENZ CHARAKTERISTIKA:
 (A) LÄNGE: 42 Basenpaare
 (B) ART: Nukleinsäure
 (C) STRANGFORM: Einzel
 (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: synthetische DNA

55

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 26:

ATGCTCGAGA ACTGCAGCAT CGATTCTGGG TCCCCCTCTG AG

42

(2) INFORMATION ZU SEQ ID NO: 27:

5 (i) SEQUENZ CHARAKTERISTIKA:
 (A) LÄNGE: 40 Basenpaare
 (B) ART: Nukleinsäure
 (C) STRANGFORM: Einzel
 (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: synthetische DNA

10 (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 27:

CGAGAACTGC AGCATCGATG CTGCTCCCAA GAACCCAAGG

40

15

(2) INFORMATION ZU SEQ ID NO: 28:

20 (i) SEQUENZ CHARAKTERISTIKA:
 (A) LÄNGE: 21 Basenpaare
 (B) ART: Nukleinsäure
 (C) STRANGFORM: Einzel
 (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: synthetische DNA

25 (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 28:

GGAGCTGCTT GATGCCAG A

21

30

(2) INFORMATION ZU SEQ ID NO: 29:

35 (i) SEQUENZ CHARAKTERISTIKA:
 (A) LÄNGE: 22 Basenpaare
 (B) ART: Nukleinsäure
 (C) STRANGFORM: Einzel
 (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: synthetische DNA

40 (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 29:

TGATGACAGC ATGTCAGGGA GT

22

45

(2) INFORMATION ZU SEQ ID NO: 30:

50 (i) SEQUENZ CHARAKTERISTIKA:
 (A) LÄNGE: 25 Basenpaare
 (B) ART: Nukleinsäure
 (C) STRANGFORM: Einzel
 (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: synthetische DNA

55 (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 30:

GCTGACATTT ATCACAGCTG GCTAC

25

(2) INFORMATION ZU SEQ ID NO: 31:

5 (i) SEQUENZ CHARAKTERISTIKA:
 (A) LÄNGE: 27 Basenpaare
 (B) ART: Nukleinsäure
 (C) STRANGFORM: Einzel
 (D) TOPOLOGIE: linear

10 (ii) ART DES MOLEKÜLS: synthetische DNA

15 (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 31:

TATCACCTAG AACCTTAAAT GCATGGG

27

(2) INFORMATION ZU SEQ ID NO: 32:

20 (i) SEQUENZ CHARAKTERISTIKA:
 (A) LÄNGE: 22 Basenpaare
 (B) ART: Nukleinsäure
 (C) STRANGFORM: Einzel
 (D) TOPOLOGIE: linear

25 (ii) ART DES MOLEKÜLS: synthetische DNA

30 (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 32:

AGTCCCTGAC ATGCTGTCAT CA

22

(2) INFORMATION ZU SEQ ID NO: 33:

35 (i) SEQUENZ CHARAKTERISTIKA:
 (A) LÄNGE: 22 Basenpaare
 (B) ART: Nukleinsäure
 (C) STRANGFORM: Einzel
 (D) TOPOLOGIE: linear

40 (ii) ART DES MOLEKÜLS: synthetische DNA

45 (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 33:

GTGGAGGGGA ATTTTCTAC TG

22

(2) INFORMATION ZU SEQ ID NO: 34:

50 (i) SEQUENZ CHARAKTERISTIKA:
 (A) LÄNGE: 24 Basenpaare
 (B) ART: Nukleinsäure
 (C) STRANGFORM: Einzel
 (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: synthetische DNA

55 (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 34:

CCTGCTGCTC CCAAGAACCC AAGG

24

(2) INFORMATION ZU SEQ ID NO: 35:

5 (i) SEQUENZ CHARAKTERISTIKA:
 (A) LÄNGE: 20 Basenpaare
 (B) ART: Nukleinsäure
 (C) STRANGFORM: Einzel
 (D) TOPOLOGIE: linear

10 (ii) ART DES MOLEKÜLS: Primer

15 (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 35:

AGCAGCAGGA AGCACTATGG

20

15 (2) INFORMATION ZU SEQ ID NO: 36:

20 (i) SEQUENZ CHARAKTERISTIKA:
 (A) LÄNGE: 20 Basenpaare
 (B) ART: Nukleinsäure
 (C) STRANGFORM: Einzel
 (D) TOPOLOGIE: linear

25 (ii) ART DES MOLEKÜLS: Primer

25 (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 36:

GAGTTTTCCA GAGCAACCCC

20

30 (2) INFORMATION ZU SEQ ID NO: 37:

35 (i) SEQUENZ CHARAKTERISTIKA:
 (A) LÄNGE: 195 Basenpaare
 (B) ART: Nukleinsäure
 (C) STRANGFORM: Einzel
 (D) TOPOLOGIE: linear

40 (ii) ART DES MOLEKÜLS: Genom-DNA

40 (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 37:

GCGCACCGGC AACAGCGCTG ACGGTACCGA CCCACAGTGT ACTGAAGGGT ATAGTGCAAC

60

AGCAGGACAA CCTGCTGAGA GCGATACAGG CCCAGCAACA CTTGCTGAGG TTATCTGTAT

120

45 GGGTATTAG ACAACTCCGA GCTCGCTGC AAGCCTTAGA AACCCCTTATA CAGAATCAGC

180

AACGCCTAAA CCTAT

195

50

55

(2) INFORMATION ZU SEQ ID NO: 38:

(i) SEQUENZ CHARAKTERISTIKA:
(A) LÄNGE: 195 Basenpaare
(B) ART: Nukleinsäure
(C) STRANGFORM: Einzel
(D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Genom-DNA

10 (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 38:

CGCGTCGCCG	TTGTCGAC	TGCCATGCCT	GGGTGTCACA	TGACTTCCA	TATCACGTTG	60
TCGTCCTGTT	GGACGACTCT	CGCTATGTCC	GGGTCGTTGT	GAACGACTCC	AATAGACATA	120
CCCCATAATC	TGTTGAGGCT	CGAGCGGACG	TTCGGAATCT	TTGGGAATAT	GTCTTAGTCG	180
TTGCGGATT	GGATA					195

20 (2) INFORMATION ZU SEQ ID NO: 39:

(i) SEQUENZ CHARAKTERISTIKA:
(A) LÄNGE: 64 Aminosäuren
(B) ART: Aminosäure
(C) STRANGFORM: Einzel
(D) TOPOLOGIE: linear

31 (ii) ART DES MOLEKÜLS: Protein

(v) ART DES FRAGMENTS: inneres

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 39:

35 Ala Ala Ala Thr Ala Leu Thr Val Arg Thr His Ser Val Leu Lys Gly
1 5 10 15

Ile Val Gln Gln Gln Asp Asn Leu Leu Arg Ala Ile Gln Ala Gln Gln
20 25 30

His Leu Leu Arg Leu Ser Val Trp Gly Ile Arg Gln Leu Arg Ala Arg
35 40 45

35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100

45 Leu Gin Ala Leu Glu Thr Leu Ile Ser Asp Ser Val Arg 100
50 55 60

50

55

(2) INFORMATION ZU SEQ ID NO: 40:

5 (i) SEQUENZ CHARAKTERISTIKA:
 (A) LÄNGE: 22 Basenpaare
 (B) ART: Nukleinsäure
 (C) STRANGFORM: Einzel
 (D) TOPOLOGIE: linear

10 (ii) ART DES MOLEKÜLS: Primer

15 (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 40:

CAGAACATCAGC AACGCCTAAA CC

22

15 (2) INFORMATION ZU SEQ ID NO: 41:

20 (i) SEQUENZ CHARAKTERISTIKA:
 (A) LÄNGE: 21 Basenpaare
 (B) ART: Nukleinsäure
 (C) STRANGFORM: Einzel
 (D) TOPOLOGIE: linear

25 (ii) ART DES MOLEKÜLS: Primer

26 (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 41:

GCCCTGTCTT ATTCTTCTAG G

21

30 (2) INFORMATION ZU SEQ ID NO: 42:

35 (i) SEQUENZ CHARAKTERISTIKA:
 (A) LÄNGE: 20 Basenpaare
 (B) ART: Nukleinsäure
 (C) STRANGFORM: Einzel
 (D) TOPOLOGIE: linear

40 (ii) ART DES MOLEKÜLS: Primer

45 (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 42:

GCCTGCAAGC CTTAGAAACC

20

45 (2) INFORMATION ZU SEQ ID NO: 43:

50 (i) SEQUENZ CHARAKTERISTIKA:
 (A) LÄNGE: 23 Basenpaare
 (B) ART: Nukleinsäure
 (C) STRANGFORM: Einzel
 (D) TOPOLOGIE: linear

55 (ii) ART DES MOLEKÜLS: Primer

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 43:

GCACTATAACC CTTCAAGTACA CTG

23

(2) INFORMATION ZU SEQ ID NO: 44:

5 (i) SEQUENZ CHARAKTERISTIKA:
 (A) LÄNGE: 1057 Basenpaare
 (B) ART: Nukleinsäure
 (C) STRANGFORM: Einzel
 (D) TOPOLOGIE: linear

10 (ii) ART DES MOLEKÜLS: Genom-DNA

15 (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 44:

AAATGTCAAG ACCAATAATA AACATTACACA CCCCTCACAG GGAAAAAAGA GCAGTAGGAT	60
TGGGAATGCT ATTCTTGGGG GTGCTAAGTG CAGCAGGTAG CACTATGGC GCAGCGGCAA	120
CAGCGCTGAC GGTACGGACC CACAGTGAC TGAAGGGTAT AGTGCACAG CAGGACAACC	180
TGCTGAGAGC GATACAGGCC CAGCAACACT TGCTGAGGTT ATCTGTATGG GGTATTAGAC	240
AACTCCGAGC TCGCCTGCAA GCCTTAGAAA CCCTTATACA GAATCAGCAA CGCCTAAACC	300
TATGGGGCTG TAAAGGAAA CTAATCTGTT ACACATCAGT AAAATGGAAC ACATCATGGT	360
CAGGAGGATA TAATGATGAC AGTATTTGGG ACAACCTTAC ATGGCAGCAA TGGGACCAAC	420
ACATAAACAA TGTAAAGCTCC ATTATATATG ATGAAATAACA AGCAGCACAA GACCAACAGG	480
AAAAGAATGT AAAAGCATTG TTGGAGCTAG ATGAATGGC CTCTCTTGG AATTGGTTG	540
ACATAACTAA ATGGTTGTGG TATATAAAAA TAGCTATAAT CATACTGGGA GCACTAATAG	600
GTATAAGAGT TATCATGATA GTACTTAATC TAGTGAAGAA CATTAGGCAG GGATATCAAC	660
CCCTCTCGTT GCAGATCCCT GTCCCACACC GGCAGGAAGC AGAAACGCCA GGAAGAACAG	720
GAGAAGAAGG TGGAGAAGGA GACAGGCCA AGTGGACAGC CTTGCCACCA GGATTCTTGC	780
AACAGTTGTA CACGGATCTC AGGACAATAA TCTTGTGGAC TTACCACTC TTGAGCAACT	840
35 TAATATCAGG GATCCGGAGG CTGATCGACT ACCTGGGACT GGGACTGTGG ATCCTGGGAC	900
AAAAGACAAT TGAAGCTTGT AGACTTTGTG GAGCTGTAAT GCAATATTGG CTACAAGAAT	960
TGAAAAATAG TGCTACAAAC CTGCTTGATA CTATTGAGT GTCAAGTTGCC AATTGGACTG	1020
40 ACGGCATCAT CTTAGGTCTA CAAAGAATAG GACAAGG	1057

45

50

55

(2) INFORMATION ZU SEQ ID NO: 45:

5	(i) SEQUENZ CHARAKTERISTIKA:	
	(A) LÄNGE: 1057 Basenpaare	
	(B) ART: Nukleinsäure	
	(C) STRANGFORM: Einzel	
	(D) TOPOLOGIE: linear	
10	(ii) ART DES MOLEKÜLS: Genom-DNA	
15	(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 45:	
15	TTTACAGTTC TGGTTATTAT TTGTAAGTGT GGGGAGTGTC CCTTTTTCT CGTCATCCTA	60
20	ACCCCTACGA TAAGAACCCC CACGATTACAC GTCGTCCATC GTGATACCCG CGTCGCCGTT	120
25	GTCGCGACTG CCATGCCTGG GTGTCACATG ACTTCCCATA TCACGTTGTC GTCCCTGTTGG	180
30	ACGACTCTCG CTATGTCCGG GTCGTTGTGA ACGACTCCAA TAGACATACC CCATAATCTG	240
35	TTGAGGCTCG AGCGGACGTT CGGAATCTT GGGAAATATGT CTTAGTCGTT GCGGATTTGG	300
40	ATACCCCGAC ATTTCTTTT GATTAGACAA TGTGTAGTCA TTTTACCTTG TGTAGTACCA	360
45	GTCCTCCTAT ATTACTACTG TCATAAACCC TGTTGGAATG TACCGTCGTT ACCCTGGTTG	420
50	TGTATTGTT ACATTCGAGG TAATATATAC TACTTATGT CGTCGTGTT CTGGTTGTCC	480
55	TTTTCTTACA TTTTCGTAAC AACCTCGATC TACTTACCCG GAGAGAAACC TTAACCAAAC	540
60	TGTATTGATT TACCAACACC ATATATTTT ATCGATATTA GTATCACCCCT CGTGATTATC	600
65	CATATTCTCA ATAGTACTAT CATGAATTAG ATCACTCTT GAAATCCGTC CCTATAGTTG	660
70	GGGAGAGCAA CGTCTAGGGA CAGGGTGTGG CCGTCCTTCG TCTTGTGGT CCTTCTTGTG	720
75	CTCTCTTCC ACCTCTTCCT CTGTCGGGT TCACCTGTCG GAAACGGTGGT CCTAAGAACG	780
80	TTGTCAACAT GTGCCTAGAG TCCTGTTATT AGAACACCTG AATGGTGGAG AACTCGTTGA	840
85	ATTATAGTCC CTAGGCCTCC GACTAGCTGA TGGACCCCTGA CCCTGACACC TAGGACCCCTG	900
90	TTTTCTGTTA ACTTCGAACA TCTGAAACAC CTCGACATTA CGTTATAACC GATGTTCTTA	960
95	ACTTTTTATC ACGATGTTG GACGAACAT GATAACGTCA CAGTCACCGG TTAACCTGAC	1020
100	TGCCGTAGTA GAATCCAGAT GTTTCTTATC CTGTTCC	1057

45

50

55

(2) INFORMATION ZU SEQ ID NO: 46:

5 (i) SEQUENZ CHARAKTERISTIKA:
 (A) LÄNGE: 351 Aminosäuren
 (B) ART: Aminosäure
 (C) STRANGFORM: Einzel
 (D) TOPOLOGIE: linear

10 (ii) ART DES MOLEKÜLS: Protein

15 (v) ART DES FRAGMENTS: inneres

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 46:

Met Ser Arg Pro Ile Ile Asn Ile His Thr Pro His Arg Glu Lys Arg
 1 5 10 15

15 Ala Val Gly Leu Gly Met Leu Phe Leu Gly Val Leu Ser Ala Ala Gly
 20 25 30

Ser Thr Met Gly Ala Ala Ala Thr Ala Leu Thr Val Arg Thr His Ser
 35 40 45

20 Val Leu Lys Gly Ile Val Gln Gln Asp Asn Leu Leu Arg Ala Ile
 50 55 60

Gln Ala Gln Gln His Leu Leu Arg Leu Ser Val Trp Gly Ile Arg Gln
 65 70 75 80

25 Leu Arg Ala Arg Leu Gln Ala Leu Glu Thr Leu Ile Gln Asn Gln Gln
 85 90 95

Arg Leu Asn Leu Trp Gly Cys Lys Gly Lys Leu Ile Cys Tyr Thr Ser
 100 105 110

30 Val Lys Trp Asn Thr Ser Trp Ser Gly Gly Tyr Asn Asp Asp Ser Ile
 115 120 125

Trp Asp Asn Leu Thr Trp Gln Gln Trp Asp Gln His Ile Asn Asn Val
 130 135 140

35 Ser Ser Ile Ile Tyr Asp Glu Ile Gln Ala Ala Gln Asp Gln Gln Glu
 145 150 155 160

Lys Asn Val Lys Ala Leu Leu Glu Leu Asp Glu Trp Ala Ser Leu Trp
 165 170 175

40 Asn Trp Phe Asp Ile Thr Lys Trp Leu Trp Tyr Ile Lys Ile Ala Ile
 180 185 190

Ile Ile Val Gly Ala Leu Ile Gly Ile Arg Val Ile Met Ile Val Leu
 195 200 205

45 Asn Leu Val Lys Asn Ile Arg Gln Gly Tyr Gln Pro Leu Ser Leu Gln
 210 215 220

Ile Pro Val Pro His Arg Gln Glu Ala Glu Thr Pro Gly Arg Thr Gly
 225 230 235 240

50 Glu Glu Gly Gly Glu Gly Asp Arg Pro Lys Trp Thr Ala Leu Pro Pro
 245 250 255

Gly Phe Leu Gln Gln Leu Tyr Thr Asp Leu Arg Thr Ile Ile Leu Trp
 260 265 270

15 Thr Tyr His Leu Leu Ser Asn Leu Ile Ser Gly Ile Arg Arg Leu Ile
 275 280 285

5 Asp Tyr Leu Gly Leu Gly Leu Trp Ile Leu Gly Gln Lys Thr Ile Glu
 290 295 300

Ala Cys Arg Leu Cys Gly Ala Val Met Gln Tyr Trp Leu Gln Glu Leu
 305 310 315 320

10 Lys Asn Ser Ala Thr Asn Leu Leu Asp Thr Ile Ala Val Ser Val Ala
 325 330 335

Asn Trp Thr Asp Gly Ile Ile Leu Gly Leu Gln Arg Ile Gly Gln
 340 345 350

15

(2) INFORMATION ZU SEQ ID NO: 47:

20 (i) SEQUENZ CHARAKTERISTIKA:
 (A) LÄNGE: 22 Basenpaare
 (B) ART: Nukleinsäure
 (C) STRANGFORM: Einzel
 (D) TOPOLOGIE: linear

25 (ii) ART DES MOLEKÜLS: Primer

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 47:

30 CTAGCAGTGG CGCCCGAACCA GG

22

(2) INFORMATION FOR SEQ ID NO: 48:

35 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 23 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

40 (ii) MOLECULE TYPE: Primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 48:

45 AATGAGGAAG CUGGAGAUTG GGA

23

50

55

(2) INFORMATION FOR SEQ ID NO: 49:

5 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 23 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

10 (ii) MOLECULE TYPE: Primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 49:

15 TCCCAUTCTG CUGCTTCCTC ATT

23

20 (2) INFORMATION ZU SEQ ID NO: 50:

(i) SEQUENZ CHARAKTERISTIKA:
 (A) LÄNGE: 26 Basenpaare
 (B) ART: Nukleinsäure
 25 (C) STRANGFORM: Einzel
 (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Primer

30 (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 50:

CCAAGGGGAA GTGACATAGC AGGAAC

26

35 (2) INFORMATION ZU SEQ ID NO: 51:

(i) SEQUENZ CHARAKTERISTIKA:
 (A) LÄNGE: 21 Basenpaare
 (B) ART: Nukleinsäure
 40 (C) STRANGFORM: Einzel
 (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Primer

45 (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 51:

CGTTGTCAG AATTCAAACC C

21

50

55

(2) INFORMATION ZU SEQ ID NO: 52:

5 (i) SEQUENZ CHARAKTERISTIKA:
 (A) LÄNGE: 21 Basenpaare
 (B) ART: Nukleinsäure
 (C) STRANGFORM: Einzel
 (D) TOPOLOGIE: linear

10 (ii) ART DES MOLEKÜLS: Primer

15 (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 52:

TCCCTAAAAA ATTAGCCTGT C

21

(2) INFORMATION ZU SEQ ID NO: 53:

20 (i) SEQUENZ CHARAKTERISTIKA:
 (A) LÄNGE: 19 Basenpaare
 (B) ART: Nukleinsäure
 (C) STRANGFORM: Einzel
 (D) TOPOLOGIE: linear

25 (ii) ART DES MOLEKÜLS: Genom-DNA

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 53:

AAACCTCCAA TTCCCCCTA

19

(2) INFORMATION ZU SEQ ID NO: 54:

30 (i) SEQUENZ CHARAKTERISTIKA:
 (A) LÄNGE: 39 Aminosäuren
 (B) ART: Aminosäure
 (C) STRANGFORM: Einzel
 (D) TOPOLOGIE: linear

35 (ii) ART DES MOLEKÜLS: Protein

(v) ART DES FRAGMENTS: inneres

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 54:

40 Cys Ile Arg Glu Gly Ile Ala Glu Val Gln Asp Ile Tyr Thr Gly Pro
 1 5 10 15

45 Met Arg Trp Arg Ser Met Thr Leu Lys Arg Ser Asn Asn Thr Ser Pro
 20 25 30

50 Arg Ser Arg Val Ala Tyr Cys
 35

55

(2) INFORMATION ZU SEQ ID NO: 55:

5 (i) SEQUENZ CHARAKTERISTIKA:
(A) LÄNGE: 41 Aminosäuren
(B) ART: Aminosäure
(C) STRANGFORM: Einzel
(D) TOPOLOGIE: linear

10 (ii) ART DES MOLEKÜLS: Protein

(v) ART DES FRAGMENTS: inneres

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 55:

15 Cys Ile Arg Glu Gly Ile Ala Glu Val Gln Asp Leu His Thr Gly Pro
1 5 10 15
Leu Arg Trp Arg Ser Met Thr Leu Lys Lys Ser Ser Asn Ser His Thr
20 25 30
20 Gln Pro Arg Ser Lys Val Ala Tyr Cys
35 40

25

30

35

40

45

50

55

(2) INFORMATION ZU SEQ ID NO: 56:

5 (i) SEQUENZ CHARAKTERISTIKA:
 (A) LÄNGE: 9793 Basenpaare
 (B) ART: Nukleinsäure
 (C) STRANGFORM: Einzel
 (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Genom-DNA

10	(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 56:	
	CTGGATGGGT TAATTTACTC CCATAAGAGA GCAGAAATCC TGGATCTCTG GATATATCAC	60
15	ACTCAGGGAT TCTTCCCTGA TTGGCAGTGT TACACACCGG GACCAGGACC TAGATTCCCA	120
	CTGACATTG GATGGTGTGTT TAAACTGGTA CCAGTGTCA GAGAGAGAGC AGAGAGACTG	180
	GGTAATACAA ATGAAGATGC TAGTCTTCTA CATCCAGCTT GTAATCATGG AGCTGAGGAT	240
20	GCACACGGGG AGATACTAAA ATGGCAGTTT GATAGATCAT TAGGCTTAAC ACATATAGCC	300
	CTGCAAAAGC ACCCAGAGCT CTTCCCCAAG TAACTGACAC TGCGGGACTT TCCAGACTGC	360
	TGACACTGCG GGGACTTTCC AGCGTGGGAG GGATAAGGGG CGGTTGGGG AGTGGCTAAC	420
25	CCTCAGATGC TGCACTATAAG CAGCTGCTT CCGCTTGTAC CGGGTCTTAG TTAGAGGACC	480
	AGGTCTGAGC CCGGGAGCTC CCTGGCCTCT AGCTGAACCC GCTGCTTAAC GCTCAATAAA	540
	GCTTGCCTTG AGTGAGAACG AGTGTGTGCT CATCTGTTCA ACCCTGGTGT CTAGAGATCC	600
30	CTCAGATCAC TTAGACTGAA GCAGAAAATC TCTAGCAGTG GCGCCCGAAC AGGGACCGA	660
	AAAGTGAAGT GGAACCAGGG AAGAAAACCT CCGACGCAAC GGGCTGGCT TAGCGGAGTG	720
	CACCTGCTAA GAGGCGAGAG GAACTCACA GAGGGTGAGT AAATTTGCTG GCGGTGGCCA	780
	GACCTAGGGG AAGGGCGAAG TCCCTAGGGG AGGAAGATGG GTGCGAGAGC GTCTGTGTTG	840
35	ACAGGGAGTA AATTGGATGC ATGGGAACGA ATTAGTTAA GGCCAGGATC TAAAAAGGCA	900
	TATAGGCTAA AACATTTAGT ATGGGCAAGC AGGGAGCTGG AAAGATACGC ATGTAATCCT	960
	GGTCTATTAG AAACTGCAGA AGGTACTGAG CAACTGCTAC AGCAGTTAGA GCCAGCTCTC	1020
40	AAGACAGGGT CAGAGGACCT GAAATCTCTC TGGAAACGCAA TAGCAGTACT CTGGTGGCTT	1080
	CACAACAGAT TTGACATCCG AGATACACAG CAGGAATAC AAAAGTTAAA GGAAGTAATG	1140
	GCAAGCAGGA AGTCTGCAGA GGCGCTAAG GAAGAAACAA GCCCTAGGCA GACAAGTCAN	1200
45	AATTACCCCTA TAGTAACAAA TGCACAGGGAA CAAATGGTAC ATCAAGCCAT CTCCCCCAGG	1260
	ACTTTAAATG CATGGTAAA GGCAGTAGAA GAGAAGGCCT TAAACCCCTGA AATTATTCCCT	1320
	ATGTTTATGG CATTATCAGA AGGGCTGTC CCCTATGATA TCAATACCAT GCTGAATGCC	1380
	ATAGGGGGAC ACCAAGGGGC TTTACAAGTG TTGAAGGAAG TAATCAATGA GGAAGCAGCA	1440
50	GAATGGGATA GAACTCATCC ACCAGCAATG GGGCCGTTAC CACCAAGGGCA GATAAGGGAA	1500
	CCAACAGGAA GTGACATTGC TGGAAACAATC AGCACACAGC AAGAGCAAAT TATATGGACT	1560

	ACTAGAGGGG CTAAC TCTAT CCCAGTAGGA GACATCTATA GAAAATGGAT AGTGCTAGGA	1620
5	CTAAACAAAA TGGTAAAAT GTACAGTCCA GTGAGCATCT TAGATATTAG GCAGGGACCA	1680
	AAAGAACCAT TCAGAGATTA TGTAGATCGG TTTACAAA CATTAGACC TGAGCAAGCT	1740
	ACTCAAGAAG TAAAGAATTG GATGACAGAA ACCTTGCTTG TTCAGAATTG AAACCCAGAT	1800
	TGTAAACAAA TTCTGAAAGC ATTAGGACCA GAAGCTACTT TAGAAGAAAT GATGGTAGCC	1860
10	TGTCAAGGAG TAGGAGGGCC AACTCACAAG GCAAAATAC TAGCAGAAGC AATGGCTTCT	1920
	GCCCAGCAAG ATTTAAAAGG AGGATAACACA GCAGTATTCA TGCAAAGAGG GCAGAATCCA	1980
	AATAGAAAAG GGCCCATAAA ATGCTTCAAT TGTGAAAAG AGGGACATAT AGCAAAAC	2040
	TGTGAGGAC CTAGAAAAG GGGTTGCTGG AAATGTGGAC AGGAAGGTCA CCAAATGAAA	2100
15	GATTGAAAAA ATGGAAGACA GGCRAATTTC TTAGGGAAAGT ACTGGCTCC GGGGGGCACG	2160
	AGGCCAGGCA ATTATGTGCA GAAACAAGTG TCCCCATCAG CCCCACCAAT GGAGGAGGCA	2220
	GTGAAGGAAC AAGAGAATCA GAGTCAGAAG GGGGATCAGG AAGAGCTGA CCCATTGCC	2280
20	TCCCTCAAAT CCCTCTTGG GACAGACCA TAGTCACAGC AAAGGTTGGG GGTCACTAT	2340
	GTGAGGCTTT ACTGGATACA GGGGAGATG ATACAGTATT AAATAACATA CAATTAGAAG	2400
	GAAGATGGAC ACCAAAATG ATAGGGGTA TAGGAGGCTT TATAAAAGTA AAAGAGTATA	2460
25	ACAATGTGAC AGTAGAAGTA CAAGGAAAGG AAGTACAGGG AACAGTATTG GTGGGACCTA	2520
	CTCCCTGTTAA TATTCTTGGG AGAACATAT TGACAGGATT AGGATGTACA CTAAATTCC	2580
	CTATAAGTCC CATAGCCCCA GTGCCAGTAA AGCTAAAACC AGGAATGGAT GGACCAAAAG	2640
	TAAAACAATG GCCCCTATCT AGAGAGAAAA TAGAAGCACT AACTGCAATA TGTCAGAAA	2700
30	TGGAACAGGA AGGAAAATC TCAAGAATAG GACCTGAAAA TCCTTATAAT ACACCTATT	2760
	TTGCTATAAA AAAGAAAGAT AGCCTAAGT GGAGAAAATT GGTAGACTTC AGAGAATTAA	2820
	ATAAAAGAAC ACAAGATTC TGGGAGGTGC AATTAGGTAT TCCACATCCA GGGGGTTAA	2880
35	AGCAAAGGCA ATCTGTTACA GTCTTAGATG TAGGAGATGC TTATTCCTCA TGCCCTTCT	2940
	ATCCAGACTT TAGAAAATAC ACTGCCTTCA CTATTCTAG TGTGAACTAT GAGACCCAG	3000
	GAGTAAGATA CCAGTACAAT GTCCCTCCGC AAGGGTGGAA AGGTTCACCA GCCATATTTC	3060
40	AGAGTTCAAT GACAAAGATT CTAGATCCAT TTAGAAAAG CAACCCAGAA GTAGAAATT	3120
	ATCAGTACAT AGATGACTTA TATGTAGGAT CAGATTACC ATTGGCAGAA CATAGAAAGA	3180
	GGGTGAAATT GCTTAGGGAA CATTATATC AGTGGGGATT TACTACCCCT GATAAAAGC	3240
	ATCAGAAGGA ACCTCCCTT TTATGGATGG GATATGAGCT CCACCCAGAC AAGTGGACAG	3300
45	TACAGCCCAT CCAATTGCCT GACAAAGAAG TGTGGACAGT AAATGATATA CAAAAATTAG	3360
	TAGGAAAATT AAATTGGGCA AGTCAAATCT ATCAAGGAAT TAGAGTAAA GAATTGTGCA	3420
	AGTTAACAG AGGAACCAA TCATTGACAG AGGTAGTACC TTTAAGTAAA GAGGCAGAAC	3480
50	TAGAATTAGA AGAAAACAGA GAAAAGCTAA AAGAGCCAGT ACATGGAGTA TATTACCGC	3540
	CTGACAAAGA CTTGTGGGTT AGTATTCAAG AGCATGGAGA AGGGCAATGG ACTTACCAAGG	3600

5	TATATCAGGA TGAACATAAG AACCTAAAAA CAGGAAAATA TGCTAGGCAA AAGGCCTCCC	3660
	ACACAAATGA TATAAGACAA TTGGCAGAAG TAGTCCAGAA GGTGTCTCAA GAAGCTATAG	3720
	TTATATGGGG GAAATTACCT AAATTCAAGC TGCCAGTTAC TAGAGAAACT TGGGAAACTT	3780
	GGTGGGCAGA ATATTGGCAG GCCACCTGGA TTCTGAATG GGAATTGTC AGCACACCCC	3840
10	CATTGATCAA ATTATGGTAC CAGTTAGAAA CAGAACCTAT TGTAGGGCA GAAACCTTTT	3900
	ATGTAGATGG AGCAGCTAAT AGGAATACAA AACTAGGAAA GGCGGGATAT GTTACAGAAC	3960
	AAGGAAAACA GAACATAATA AAGTTAGAAG AGACAACCAA TCAAAAGGCT GAAATATGG	4020
	CTGTATTAAT AGCCTTGCAG GATTCCAAGG AGCAAGTAAA CATACTAAC AACTCACAAT	4080
15	ATGTATGGG CATCATATCC TCCCACCAA CACAGAGTGA CTCCCCTATA GTTCAGCAGA	4140
	TAATAGAGGA ACTAACAAA AAGGAACGAG TGTATCTTAC ATGGGTTCT GCTCACAAAG	4200
	GCATAGGAGG AAATGAAAAA ATAGATAAT TAGTAAGCAA AGACATTAGA AGAGTCCTGT	4260
	TCCTGGAAGG AATAGATCAG GCACAAAGAAG ATCATGAAA ATATCATACT AATTGGAGAG	4320
20	CATTAGCTAG TGACTTTGGA TTACCAACCA TAGTAGCCAA GGAATCATT GCTAGTTGTC	4380
	CTAAATGCCA TATAAAAGGG GAAGCAACGC ATGGTCAAGT AGACTACAGC CCAGAGATAT	4440
	GGCAAATGGA TTGTACACAT TTAGAAGGCA AAATCATAAT AGTTGCTGTC CATGTAGCAA	4500
	GTGACTTTAT AGAACAGAG GTGATACCAAG CAGAACAGG ACAGGAAACT GCCTATTCC	4560
25	TGTTAAAATT AGCAGCAAGA TGGCCTGTCA AAGTAATACA TACAGACAAT GGACCTAATT	4620
	TTACAAGTGC AGCCATGAAA CCTGCATGTT GGTGGACAGG CATAACAAT GAGTTGGGA	4680
	TACCATATAA TCCACAAAGT CAAGGAGTAG TAGAACCCAT GAATAAGAA TTAAAATCTA	4740
30	TTATACAGCA GGTGAGGGAC CAAGCAGAGC ATTTAAAAC AGCAGTACAA ATGGCAGTCT	4800
	TTGTTCACAA TTTTAAAAGA AAAGGGGGGA TTGGGGGGTA CACTGCAGGG GAGAGACTAA	4860
	TAGACATACT AGCATCACAA ATACAAACAA CAGAACCTACA AAAACAAATT TTAAAATCA	4920
35	ACAATTTTCG GGTCTATTAC AGAGATAGCA GAGACCTAT TTGGAAAGGA CGGGCACAC	4980
	TCCITGGAA AGGTGAGGGG GCAGTGTCA TACAGATAA AGGAGACATT AAAGTGGTAC	5040
	CAAGAAGAAA GGCAAAATA ATCAGAGATT ATGGAAAACA GATGGCAGGT ACTGATAGTA	5100
40	TGGCAAATAG ACAGACAGAA AGTGAAGCA TGGAACAGCC TGGTAAATA CCATAAAATAC	5160
	ATGTCTAAGA AGGCCCGCAA CTGGCGTTAT AGGCATCATT ATGAATCCAG GAATCCAAA	5220
	GTCAGTTCGG CGGTGTATAT TCCAGTAGCA GAAGCTGATA TAGTGGTCAC CACATATTGG	5280
	GGATTAATGC CAGGGGAAAG AGAGGAACAC TTGGGACATG GGGTTAGTAT AGAATGGCAA	5340
45	TACAAGGAGT ATAAAACACA GATTGATCCT GAAACAGCAG ACAGGATGAT ACATCTGCAT	5400
	TATTTCACAT GTTTTACAGA ATCAGCAATC AGGAAGGCCA TTCTAGGGCA GAGAGTGCTG	5460
	ACCAAGTGTG AATACACCTGGC AGGACATAGT CAGGTAGGGA CACTACAATT CTTAGCCTTG	5520
50	AAAGCAGTAG TGAAAGTAAA AAGAAATAAG CCTCCCTAC CCAGTGTCCA GAGATTAACA	5580

5	GAAGATAGAT GGAACAAGCC CTGGAAAATC AGGGACCAGC TAGGGAGCCA TTCAATGAAT GGACACTAGA GCTCCTGGAA GAGCTGAAAG AAGAAGCAGT AAGACATTC CCTGCCCTT GGTTACAAGC CTGTGGGCAG TACATTTATG AGACTTATGG AGACACTTGG GAAGGGAGTTA TGGCAATTAT AAGAATCTTA CAACAACATAC TGTACCCCA TTATAGAATT GGATGCCAAC ATAGTAGAAT AGGAATTCTC CCATCTAACCA CAAGAGGAAG AGGAAGAAGA AATGGATCCA GTAGATCCTG AGATGCCCTT TTGGCATCAC CCTGGGAGCA AGCCCCAAC CCCTGTAAAT ATTGCTTATT GCAAAAGATG CTGCTATCAT TGCTATGTTT GTTTCACAAA GAAGGGTTTG GGAATCTCCC ATGGCAGGAA GAAGCGAAGA AGACCAGCAG CTGCTGCAAG CTATCCAGAT ATAAAAGATC CTGTACCCAGA GCAGTAAGTA ACGCTGATGC ATCAAGAGAA CCTGCTAGCC TTAATAGCTT TAAGTGCTTT GTGTCTTATA AATGTACTTA TATGGTTGTT TAACCTTAGA ATTTATTTAG TGCAAAGAAA ACAAGATAGA AGGGAGCAGG AAATACTTGA AAGATTAAGG AGAATAAAGG AAATCAGGG TGACAGTGAC TATGAAGTA ATGAAGAAGA ACAACAGGAA GTCATGGAGC TTATACATAG CCATGGCTTT GCTAATCCCA TGTTTGAGTT ATAGTAAACA ATTGTATGCC ACAGTTTATT CTGGGTACCC TGATGGGAA GAGGCAGCAC CAGTACTATT CTGTGCTTCA GATGCTAACCC TAACAAGCAC TGAAACAGCAT AATATTTGGG CATCACAAGC CTGCGTTCT ACAGATCCCA ATCCACATGA ATTCCACTA GGCAATGTGA CAGATAACTT TGATATATGG AAAAATTACA TGGTGGACCA AATGCATGAA GACATCATTAA GTTGTGGGA ACAGAGTTA AAGCCTTGTG AGAAAATGAC TTTCTTATGT GTACAAATGA ACTGTGTAGA TCTGCAAACA AATAAAACAG GCCTATTAAA TGAGACATAA AATGAGATGA GAAATTGTAG TTTTAATGTA ACTACAGTCC TCACAGACAA AAAGGAGCAA AAACAGGCTC TATTCTATGT ATCAGATCTG AGTAAGGTTA ATGACTCAA TGCACTAAAT GGAACAAACAT ATATGTTAAC TAATTGTAAAC TCCACAAATTCA TCAAGCAGGC CTGTCGGAGG GTAAAGTTTG AGCCATTCC CATACACTAT TGTGCTCCAA CAGGATATGC CATCTTAAAG TGTAATGACA CAGACITTTAA TGGAACAGGC CTATGCCACA ATATTCAGT GTTACTGT ACACATGGCA TCAAGCCAAC AGTAAGTACT CAACTAATAC TGAAATGGAC ACTCTCTAGA GAAAAGATAA GAATTATGGG AAAAAAATATT ACAGAACATCG CAAAGAATAT CATACTAACCT CTAACACATC CTATAAACAT GACCTGCATA AGAGAAGGAA TTGCAGAGGT ACAAGATATA TATACTGGTC CAAATGAGATG GGCGAGTATG ACACTTAAA GAAGTAACAA TACATCACCAGA AGATCAAGGG TAGCTTATTG TACATATAAT AAGACTGTAT GGGAAAATGC CCTACAAACAA ACAGCTATAA GGTATTTAAA TCTTGAAAC CAAACAGAGA ATGTTACCAT AATATTCAAGC AGAACTAGTG GTGGAGATGC AGAAGTAAGC CATTACATT TTAACGTCA TGGAGAATTC TTTTATTGTA ACACATCTGG GATGTTAAC TATACTTTA TCAACTGTAC AAAGTCCGGA TGCCAGGAGA TCAAAGGGAG CAATGAGACC AATAAAAATG GTACTATACC TTGCAAGTTA AGACAGCTAG TAAGATCATG GATGAAGGGAA GAGTCGAGAA TCTATGCACC TCCCACCCCC GGCAACTTAA CATGTCATT	564C 570C 576C 5820 5880 5940 6000 6060 6120 6180 6240 6300 6360 6420 6480 6540 6600 6660 6720 6780 6840 6900 6960 7020 7080 7140 7200 7260 7320 7380 7440 7500 7560 7620
---	--	--

	CAACATAACT GGAATGATTG TACAGTTAGA TCAACCATGG AATTCCACAG GTGAAAATAC	7680
5	ACTTAGACCA GTAGGGGGAG ATATGAAGA TATATGGAGA ACTAATTTGT ACAACTACAA	7740
	AGTAGTACAG ATAAAACCTT TTAGTGTAGC ACCTACAAA ATGTCAGAC CAATAATAAA	7800
	CATTCACACC CCTCACAGGG AAAAAAGAGC AGTAGGATTG GGATGCTAT TCTTGGGGT	7860
10	GCTAAGTGC A GCAGGTAGCA CTATGGCGC AGCGGCACCA GCGCTGACGG TACGGACCC	7920
	CAGTGTACTG AAGGGTATAG TGCAACAGCA GGACAACTG CTGAGAGCGA TACAGGCCA	7980
	GCAACACTG CTGAGGTAT CTGTATGGG TATTAGACAA CTCCGAGCTC GCCTGCAAGC	8040
	CTTAGAAACC CTTATACAGA ATCAGCAACG CCTAAACCTA TGGGGCTGTA AAGGAAAATC	8100
15	AATCTGTTAC ACATCAGTAA AATGGAACAC ATCATGGTCA GGAAGATATA ATGATGACAG	8160
	TATTTGGAC AACCTTACAT GGCAGCAATG GGACCAACAC ATAAAACAATG TAAGCTCCAT	8220
	TATATATGAT GAAATACAAG CAGCACAAAGA CCAACAGGAA AAGAATGTAA AAGCATTGTT	8280
20	GGAGCTAGAT GAATGGGCCT CTCTTGGAA TTGGTTGAC ATAACTAAAT GGTGTTGGTA	8340
	TATAAAAATA GCTATAATCA TAGTGGGAGC ACTAATAGGT ATAAAGAGTTA TTATGATAAT	8400
	ACTTAATCTA GTGAAGAAC A TTAGGCAGGG ATATCAACCC CTCTCGTTGC AGATCCCTGT	8460
	CCCACACCGG CAGGAAGCAG AAACGCCAGG AAGAACAGGA GAAAAGGTG GAGAAGGAGA	8520
25	CAGGCCAAG TGGACAGCCT TGCCACCAGG ATTCTGCAA CAGTTGTACA CGGATCTCAG	8580
	GACAATAATC TTGTGGACTT ACCACCTCTT GAGCAACTTA ATATCAGGGA TCCGGAGGCT	8640
	GATCGACTAC CTGGGACTGG GACTGTGGAT CCTGGACAA AAGACAATTG AAGCTTGTAG	8700
30	ACTTTGTGGA GCTGTAATGC AATATGGCT ACAAGAATTG AAAAATAGTG CTACAAACCT	8760
	GCTTGATACT ATTGCAGTGT CAGTTGCCAA TTGGACTGAC GGCATCATCT TAGGTCTACA	8820
	AAGAATAGGA CAAGGATTCC TTCACATCCC AAGAAGAATT AGACAAGGTG CAGAAAAGAT	8880
35	CTTAGTGTAA CATGGGAAT GCATGGAGCA AAAGCAAATT TGCAAGGATGG TCAGAAGTAA	8940
	GAGATAGAAAT GAGACGATCC TCCCTGATC CTCAACAAACC ATGTGCACCT GGAGTAGGAG	9000
	CTGTCTCCAG GGAGTTAGCA ACTAGAGGGG GAATATCAAG TTCCCACACT CCTCAAAACA	9060
40	ATGCAGCCCT TGCAATTCTA GACAGCCACA AAGATGAGGA TGTAGGCTTC CCAGTAAGAC	9120
	CTCAAGTGCC TCTAAGGCCA ATGACCTTA AAGCAGCCT TGACCTCAGC TTCTTTTAA	9180
	AAGAAAAGGG AGGACTGGAT GGGTTAATT ACTCCCATAA GAGAGCAGAA ATCTGGATC	9240
	TCTGGATATA TCACACTCAG GGATTCTTCC CTGATTGGCA GTGTTACACA CCGGGACCAAG	9300
45	GACCTAGATT CCCACTGACA TTTGGATGGT TGTTAAACT GGTACCACTG TCAGCAGAAG	9360
	AGGCAGAGAG ACTGGGTAAT ACAAAATGAAG ATGCTAGTCT TCTACATCCA GCTTGTAAATC	9420
	ATGGAGCTGA GGATGCACAC GGGGAGATAC TAAAATGGCA GTTTGATAGA TCATTAGGCT	9480
50	TAACACATAT AGCCCTGCCAA AAGCACCCAG AGCTCTTCCC CAAGTAACTG ACACGTGCGGG	9540
	ACTTTCCAGA CTGCTGACAC TGCGGGACT TTCCAGCGTG GGAGGGATAA GGGCGGGTTC	9600

GGGGAGTGGC TAACCCTCA G ATGCTGCATA TAAGCAGCTG CTTTCCGCTT GTACCGGGTC	9660
TTAGTTAGAG GACCAGGTCT GAGCCCGGGA GCTCCCTGGC CTCTAGCTGA ACCCGCTGCT	9720
5 TAACGCTCAA TAAAGCTTGC CTTGAGTGTG AAGCAGTGTG TGCTCATCTG TTCAACCCCTG	9780
GTGTCTAGAG ATC	9793

10

15

20

25

31

35

4

1

4

11

(2) INFORMATION ZU SEQ ID NO: 57:

5 (i) SEQUENZ CHARAKTERISTIKA:
 (A) LÄNGE: 1733 Basenpaare
 (B) ART: Nukleinsäure
 (C) STRANGFORM: Einzel
 (D) TOPOLOGIE: linear

10 (ii) ART DES MOLEKÜLS: Genom-DNA

10 (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 57:

AAACCTCCGA CGCAACGGGC TCGGCTTAGC GGAGTCACC TGCTAAGAGG CGAGAGGAAC	60
TCACAAGAGG GTGAGTAAAT TTGCTGGCGG TGGCCAGACC TAGGGGAAGG GCGAAGTCCC	120
15 TAGGGGAGGA AGATGGGTGC GAGAGCGCT GTGTTGACAG GGAGTAAATT GGATGCATGG	180
GAACGAATTA GGTTAAGGCC AGGATCTAAA AAGGCATATA GGCTAAAACA TTTAGTATGG	240
20 GCAAGCAGGG AGCTGGAAAG ATACGCATGT AATCCTGGTC TATTAGAAAC TGCAGAAGGT	300
ACTGAGCAAC TGCTACAGCA GTTAGAGCCA GCTCTCAAGA CAGGGTCAGA GGACCTGAAA	360
TCTCTCTGGA ACGCAATAGC AGTACTCTGG TGCCTTCACA ACAGATTGTA CATCCGAGAT	420
ACACAGCAGG CAATACAAA GTTAAAGGAA GTATGGCAA GCAGGAAGTC TGCAGAGGCC	480
25 GCTAAGGAAG AAACAAGCCC TAGGCAGACA AGTCAAAATT ACCCTATAGT AAACAAATGCA	540
CAGGGACAAA TGGTACATCA AGCCATCTCC CCCAGGACTT TAAATGCATG GGTAAAGGCA	600
GTAGAAAGAGA AGGCCTTAA CCCTGAAATT ATTCCTATGT TTATGGCATT ATCAGAAGGG	660
30 GCTGTCCCCCT ATGATATCAA TACCATGCTG AATGCCATAG GGGGACACCA AGGGGCTTTA	720
CAAGTGTGA AGGAAGTAAT CAATGAGGAA GCAGCAGAAT GGGATAGAAC TCATCCACCA	780
GCAATGGGGC CGTTACCACC AGGGCAGATA AGGGAACCAA CAGGAAGTGA CATTGCTGGA	840
35 ACAACTAGCA CACAGCAAGA GCAAATTATA TGGACTACTA GAGGGGCTAA CTCTATCCCA	900
GTAGGAGACA TCTATAGAAA ATGGATAGTG CTAGGACTAA ACAAAATGGT AAAATGTAC	960
AGTCCAGTGA GCATCTTAGA TATTAGGCAG GGACCAAAAG AACCATTCAAG AGATTATGTA	1020
GATCGGTTTT ACAAAACATT AAGAGCTGAG CAAGCTACTC AAGAAGTAAA GAATTGGATG	1080
40 ACAGAAACCT TGCTTGTCA GAATTCAAC CCAGATTGTA AACAAATTCT GAAAGCATTAA	1140
GGACCCAGAAG CTACTTTAGA AGAAATGATG GTAGCCTGTC AAGGAGTAGG AGGGCCAACT	1200
CACAAGGCAA AAATACTAGC AGAAGCAATG GCTTCTGCCCA AGCAAGATTT AAAAGGAGGA	1260
45 TACACAGCAG TATTCTATGCA AAGAGGGCAG AATCCAATA GAAAAGGGCC CATAAAATGC	1320
TTCAATTGTG GAAAAGAGGG ACATATAGCA AAAAATGTC GAGCACCTAG AAAAAGGGGT	1380
TGCTGGAAAT GTGGACAGGA AGGTCACCAA ATGAAAGATT GCAAAATGG AAGACAGGCA	1440
50 AATTTTTTAG GGAAGTACTG GCCTCCGGGG GGCACGAGGC CAGGCAATTA TGTGCAGAAA	1500
CAAGTGTCCC CATCAGCCCC ACCAATGGAG GAGGCAGTGA AGGAACAAGA GAATCAGAGT	1560

CAGAAGGGGG ATCAGGAAGA GCTGTACCCA TTTGCCTCCC TCAAATCCCT CTTGGGACA	1620
GACCAATAGT CACAGCAAAG GTTGGGGGTC ATCTATGTGA GGCTTTACTG GATACAGGGG	1680
CAGATGATAAC AGTATTAAAT AACATACAAT TAGAAGGAAG ATGGACACCA AAA	1733

5

10

15

20

25

30

35

40

45

50

55

(2) INFORMATION ZU SEQ ID NO: 58:

5 (i) SEQUENZ CHARAKTERISTIKA:
 (A) LÄNGE: 1733 Basenpaare
 (B) ART: Nukleinsäure
 (C) STRANGFORM: Einzel
 (D) TOPOLOGIE: linear

10 (ii) ART DES MOLEKÜLS: Genom-DNA

10	(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 58:	
	AAACCTCAA CGCAACGGGC TCGGCTTAGC GGAGTGCACC TGCTAAGAGG CGAGAGGAAC	60
	TCACAAGAGG GTGAGTAAAT TTGCTGGCGG TGGCCAGACC TAGGGGAAGG GCGAAGTCCC	120
15	TAGGGGAGGA AGATGGGTGC GAGACGGTCT GTGTTGACAG GGAGTAAATT GGATGCATGG	180
	GAACGAATTA GGTTAAGGCC AGGATCTAAA AAGGCATATA GGCTAAAACA TTTAGTATGG	240
	GCAAGCAGGG AGCTGGAAAG ATACGCATAT AATCCTGGTC TACTAGAAAC TGCAGAAGGT	300
20	ACTGAACAAAC TGCTACAGCA GTTAGAGCCA GCTCTCAAGA CAGGGTCAGA GGACCTGAAA	360
	TCCCTCTGGA ACGCAATAGC AGTACTCTGG TGCCTTCACA ACAGATTGATGATCCGAGAT	420
	ACACAGCAGG CAATACAAA GTTAAAGGAA GTAAATGGCA GCAGGAAGTC TGCAGAGGCC	480
25	GCTAAGGAAG AAACAAGCTC AAGGCAGGCC AGTCAAAATT ACCCTATAGT AACAAATGCA	540
	CAGGGACAAA TGGTACATCA AGCCATATCC CCTAGGACTT TAAATGCATG GGTAAAGGCA	600
	GTAGAAGAAA AGGCCTTAA CCTCTGAAATT ATTCCTATGT TTATGGCATT ATCAGAAGGG	660
30	GCTGTCCCCCT ATGATATCAA TACCATGCTG AATGCCATAG GGGGACACCA AGGGGCTTTA	720
	CAAGTGTGTA AGGAAGTAAT CAATGAGGAA GCAGCAGATT GGGATAGAAC TCATCCACCA	780
	GCAATGGGGC CGTTACCACC AGGGCAGATA AGGGAACCAA CAGGAAGTGA CATTGCTGGA	840
35	ACAACCTAGCA CACAGCAAGA GCAAAATTATA TGGACTACTA GAGGGGCTAA CTCTATCCCA	900
	GTAGGAGACA TCTATAGAAA ATGGATAGTG TTAGGACTAA ACAAAATGGT AAAAATGTAC	960
	AGTCCAGTGA GCATCTTAGA TATTAAGGCAG GGACCAAAAG AACCAATTGAG AGATTATGTA	1020
40	GATCGGTTTT ACAAAACATT AAGAGCTGAG CAAGCTACTC AAGAAGTAAA GAATTGGATG	1080
	ACAGAAACCC TCGTTGTTCA GAATTCAAAC CCAGATTGTA AACAAATTCT GAAAGCATTAA	1140
	GGACCAGGGAG CTACTTTAGA AGAAAATGATG GTAGCCTGTC AAGGAGTAGG AGGGCCAACT	1200
	CACAAGGCCA AAATACTAGC AGAAGCAATG GCTTCTGCC AGCAAGATTT AAAGGGAGGA	1260
45	TACACAGCAG TATTCATGCA AAGAGGGCAG AATCCAAATA GAAAAGGGCC TATAAAATGT	1320
	TTCAATTGTC GAAAAGAGGG ACATATAGCA AAAAATGTC GAGCACCTAG AAGAAGGGGT	1380
	TACTGGAAAT GTGGACAGGA AGGTCAACCA ATGAAAGATT GCAAAATGG AAGACAGGCT	1440
50	ATTTTTTTAG GGAAGTACTG GCCTCCGGGG GGCACGAGGC CAGCCAATTAA TGTGCAGAAA	1500
	CAAGTGTCCC CATCAGCCCC ACCAATGGAG GAGGCAGTGA AGGAACAAGA GAATCAGAAT	1560

5 CAAAAGGGGG ATCAGGAAGA GCTGTACCCA TTTGCCTCCC TCAAATCCCT CTTGGGACA 1620
 GACCAATAGT CACAGCAAAG GTTGGGGGCC ATCTATGTGA GGCTTTACTG GATAACAGGGG 1680
 CAGATGATAC AGTATTAAT AACATACAAT TAGAAGGAAG ATGGACACCC AAA 1733

(2) INFORMATION ZU SEQ ID NO: 59:

10 (i) SEQUENZ CHARAKTERISTIKA:
 (A) LÄNGE: 498 Aminosäuren
 (B) ART: Aminosäure
 (C) STRANGFORM: Einzel
 (D) TOPOLOGIE: linear

15 (ii) ART DES MOLEKÜLS: Protein

(v) ART DES FRAGMENTS: inneres

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 59:

20 Met Gly Ala Arg Ala Ser Val Leu Thr Gly Ser Lys Leu Asp Ala Trp
 1 5 10 15
 Glu Arg Ile Arg Leu Arg Pro Gly Ser Lys Lys Ala Tyr Arg Leu Lys
 20 25 30
 His Leu Val Trp Ala Ser Arg Glu Leu Glu Arg Tyr Ala Cys Asn Pro
 25 35 40 45
 Gly Leu Leu Glu Thr Ala Glu Gly Thr Glu Gln Leu Leu Gln Gln Leu
 30 50 55 60
 Glu Pro Ala Leu Lys Thr Gly Ser Glu Asp Leu Lys Ser Leu Trp Asn
 35 65 70 75 80
 Ala Ile Ala Val Leu Trp Cys Val His Asn Arg Phe Asp Ile Arg Asp
 40 85 90 95
 Thr Gln Gln Ala Ile Gln Lys Leu Lys Glu Val Met Ala Ser Arg Lys
 45 100 105 110
 Ser Ala Glu Ala Ala Lys Glu Glu Thr Ser Pro Arg Gln Thr Ser Gln
 50 115 120 125
 Asn Tyr Pro Ile Val Thr Asn Ala Gln Gly Gln Met Val His Gln Ala
 55 130 135 140
 Ile Ser Pro Arg Thr Leu Asn Ala Trp Val Lys Ala Val Glu Lys
 60 145 150 155 160
 Ala Phe Asn Pro Glu Ile Ile Pro Met Phe Met Ala Leu Ser Glu Gly
 65 165 170 175
 Ala Val Pro Tyr Asp Ile Asn Thr Met Leu Asn Ala Ile Gly Gly His
 70 180 185 190
 Gln Gly Ala Leu Gln Val Leu Lys Glu Val Ile Asn Glu Ala Ala
 75 195 200 205
 Glu Trp Asp Arg Thr His Pro Pro Ala Met Gly Pro Leu Pro Pro Gly
 80 210 215 220
 Gln Ile Arg Glu Pro Thr Gly Ser Asp Ile Ala Gly Thr Thr Ser Thr
 85 225 230 235 240
 Gln Gln Glu Gln Ile Ile Trp Thr Thr Arg Gly Ala Asn Ser Ile Pro
 90 245 250 255
 Val Gly Asp Ile Tyr Arg Lys Trp Ile Val Leu Gly Leu Asn Lys Met
 95 260 265 270

Val Lys Met Tyr Ser Pro Val Ser Ile Leu Asp Ile Arg Gln Gly Pro
 275 280 285
 5 Lys Glu Pro Phe Arg Asp Tyr Val Asp Arg Phe Tyr Lys Thr Leu Arg
 290 295 300
 Ala Glu Gln Ala Thr Gln Glu Val Lys Asn Trp Met Thr Glu Thr Leu
 305 310 315 320
 10 Leu Val Gln Asn Ser Asn Pro Asp Cys Lys Gln Ile Leu Lys Ala Leu
 325 330 335
 Gly Pro Glu Ala Thr Leu Glu Glu Met Met Val Ala Cys Gln Gly Val
 15 340 345 350
 Gly Gly Pro Thr His Lys Ala Lys Ile Leu Ala Glu Ala Met Ala Ser
 355 360 365
 Ala Gln Gln Asp Leu Lys Gly Gly Tyr Thr Ala Val Phe Met Gln Arg
 20 370 375 380
 Gly Gln Asn Pro Asn Arg Lys Gly Pro Ile Lys Cys Phe Asn Cys Gly
 385 390 395 400
 25 Lys Glu Gly His Ile Ala Lys Asn Cys Arg Ala Pro Arg Lys Arg Gly
 405 410 415
 Cys Trp Lys Cys Gly Gln Glu Gly His Gln Met Lys Asp Cys Lys Asn
 420 425 430
 30 Gly Arg Gln Ala Asn Phe Leu Gly Lys Tyr Trp Pro Pro Gly Gly Thr
 435 440 445
 Arg Pro Gly Asn Tyr Val Gln Lys Gln Val Ser Pro Ser Ala Pro Pro
 450 455 460
 35 Met Glu Glu Ala Val Lys Glu Gln Glu Asn Gln Ser Gln Lys Gly Asp
 465 470 475 480
 Gln Glu Glu Leu Tyr Pro Phe Ala Ser Leu Lys Ser Leu Phe Gly Thr
 40 485 490 495
 Asp Gln

45

50

55

(2) INFORMATION ZU SEQ ID NO: 60:

5 (i) SEQUENZ CHARAKTERISTIKA:
 (A) LÄNGE: 498 Aminosäuren
 (B) ART: Aminosäure
 (C) STRANGFORM: Einzel
 (D) TOPOLOGIE: linear

10 (ii) ART DES MOLEKÜLS: Protein

(v) ART DES FRAGMENTS: inneres

15 (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 60:

Met Gly Ala Arg Arg Ser Val Leu Thr Gly Ser Lys Leu Asp Ala Trp
 1 5 10 15

Glu Arg Ile Arg Leu Arg Pro Gly Ser Lys Lys Ala Tyr Arg Leu Lys
 20 25 30

His Leu Val Trp Ala Ser Arg Glu Leu Glu Arg Tyr Ala Tyr Asn Pro
 35 40 45

Gly Leu Leu Glu Thr Ala Glu Gly Thr Glu Gln Leu Leu Gln Gln Leu
 50 55 60

Glu Pro Ala Leu Lys Thr Gly Ser Glu Asp Leu Lys Ser Leu Trp Asn
 65 70 75 80

Ala Ile Ala Val Leu Trp Cys Val His Asn Arg Phe Asp Ile Arg Asp
 85 90 95

Thr Gln Gln Ala Ile Gln Lys Leu Lys Glu Val Met Ala Ser Arg Lys
 100 105 110

Ser Ala Glu Ala Ala Lys Glu Glu Thr Ser Ser Thr Gln Ala Ser Gln
 115 120 125

Asn Tyr Pro Ile Val Thr Asn Ala Gln Gly Gln Met Val His Gln Ala
 130 135 140

Ile Ser Pro Arg Thr Leu Asn Ala Trp Val Lys Ala Val Glu Glu Lys
 145 150 155 160

Ala Phe Asn Pro Glu Ile Ile Pro Met Phe Met Ala Leu Ser Glu GLY
 165 170 175

Ala Val Pro Tyr Asp Ile Asn Thr Met Leu Asn Ala Ile Gly Gly His
 180 185 190

Gln Gly Ala Leu Gln Val Leu Lys Glu Val Ile Asn Glu Glu Ala Ala
 195 200 205

Asp Trp Asp Arg Thr His Pro Pro Ala Met Gly Pro Leu Pro Pro Gly
 210 215 220

45 Gln Ile Arg Glu Pro Thr Gly Ser Asp Ile Ala Gly Thr Thr Ser Thr
 225 230 235 240

Gln Gln Gln Ile Ile Trp Thr Thr Arg Gly Ala Asn Ser Ile Pro
 245 250 255

50 Val Gly Asp Ile Tyr Arg Lys Trp Ile Val Leu Gly Leu Asn Lys Met
 260 265 270

Val Lys Met Tyr Ser Pro Val Ser Ile Leu Asp Ile Arg Gln Gly Pro
 275 280 285
 5 Lys Glu Pro Phe Arg Asp Tyr Val Asp Arg Phe Tyr Lys Thr Leu Arg
 290 295 300
 Ala Glu Gln Ala Thr Gln Glu Val Lys Asn Trp Met Thr Glu Thr Leu
 305 310 315 320
 10 Val Val Gln Asn Ser Asn Pro Asp Cys Lys Gln Ile Leu Lys Ala Leu
 325 330 335
 Gly Pro Gly Ala Thr Leu Glu Glu Met Met Val Ala Cys Gln Gly Val
 15 340 345 350
 Gly Gly Pro Thr His Lys Ala Lys Ile Leu Ala Glu Ala Met Ala Ser
 355 360 365
 20 Ala Gln Gln Asp Leu Lys Gly Gly Tyr Thr Ala Val Phe Met Gln Arg
 370 375 380
 Gly Gln Asn Pro Asn Arg Lys Gly Pro Ile Lys Cys Phe Asn Cys Gly
 385 390 395 400
 25 Lys Glu Gly His Ile Ala Lys Asn Cys Arg Ala Pro Arg Arg Gly
 405 410 415
 Tyr Trp Lys Cys Gly Gln Glu Gly His Gln Met Lys Asp Cys Lys Asn
 420 425 430
 30 Gly Arg Gln Ala Asn Phe Leu Gly Lys Tyr Trp Pro Pro Gly Gly Thr
 435 440 445
 Arg Pro Ala Asn Tyr Val Gln Lys Gln Val Ser Pro Ser Ala Pro Pro
 35 450 455 460
 Met Glu Glu Ala Val Lys Glu Gln Glu Asn Gln Asn Gln Lys Gly Asp
 465 470 475 480
 40 Gln Glu Glu Leu Tyr Pro Phe Ala Ser Leu Lys Ser Leu Phe Gly Thr
 485 490 495
 Asp Gln

45 **Patentansprüche**

1. Immunschwäche-Virus der HIV-Gruppe oder Varianten dieses Virus, das die wesentlichen morphologischen und immunologischen Eigenschaften des bei der European Collection of Animal Cell Cultures (ECACC) unter der Nr. V 920 92 318 hinterlegten Retrovirus mit der Bezeichnung MVP-5180/91 aufweist.
- 50 2. Immunschwäche-Virus nach Anspruch 1, dadurch gekennzeichnet, daß es eine der Reversen Transkriptase entsprechende Proteinbande im Western Blot aufweist, die 3-7 Kiloton kleiner ist als die entsprechende Bande der Viren HIV-1 und/oder HIV-2.
- 55 3. Immunschwäche-Virus nach einem der Ansprüche 1 oder 2, dadurch gekennzeichnet, daß dieses Retrovirus mit einem gegen das Protein p 24 gerichteten monoklonalen Antikörper weniger Reaktivität

aufweist, bezogen auf die Reverse Transkriptase-Aktivität, als das Virus HIV-1 und mehr Aktivität, bezogen auf die Aktivität der Reversen Transkriptase, als HIV-2.

4. Immunschwäche-Virus nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, daß mit seinem Transmembranprotein gp 41 Antigen-Antikörperreaktionen gut nachweisbar sind mit Seren von aus Afrika stammenden Patienten und, daß mit dem gp 41 nur eine geringer oder keine Antigen-Antikörperreaktion mit Seren von aus Deutschland stammenden Patienten nachgewiesen werden kann.
5. Immunschwäche-Virus nach einem der obengenannten Ansprüche, dadurch gekennzeichnet, daß es eine RNA-Sequenz aufweist, die mit der RNA des hinterlegten Virus zu etwa 75 % oder mehr, bezogen auf das Gesamtgenom, homolog ist.
10. Immunschwäche-Virus nach einem der oben genannten Ansprüche, dadurch gekennzeichnet, daß es eine RNA-Sequenz aufweist, die zu der RNA-Sequenz von Tabelle 1 zu wenigstens 75 % homolog ist.
15. Immunschwäche-Virus nach einem der Ansprüche 1 bis 5, dadurch gekennzeichnet, daß es eine Nucleotid-Sequenz aufweist, die zu der Sequenz von Tabelle 3 oder Teilen davon zu wenigstens 75 % homolog ist.
20. Immunschwäche-Virus nach Anspruch 7, dadurch gekennzeichnet, daß der Teil der Sequenz wenigstens 50 Nucleotide lang ist.
25. Immunschwäche-Virus nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, daß es eine Sequenz oder Teilsequenz aufweist, die der Fig. 4 entspricht oder zu dieser Sequenz homolog ist, wobei die Unterschiede zu der in Fig. 4 angegebenen Sequenz bezogen auf die Genorte höchstens betragen: LTR: 17 %, GAG: 29 %; POL: 25 %; VIF: 31 %; ENV: 46 %; NEF: 16 %.
30. Immunschwäche-Virus nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, daß es eine Sequenz oder Teilsequenz aufweist, die der Fig. 4 entspricht oder zu dieser Sequenz homolog ist, wobei die Unterschiede zu der in Fig. 4 angegebenen Sequenz bezogen auf die Genorte höchstens betragen: LTR: 10 %; GAG: 14 %; POL: 12 %; VIF: 15 %; ENV: 22 %; NEF: 10 %.
35. cDNA, die komplementär ist zu der RNA oder Teilen davon, das bei der European Collection of Animal Cell Cultures (ECACC) unter der Nr. V 920 92 318 hinterlegten Immunschwäche-Virus MVP-5180/81 oder eines Virus gemäß einem der Ansprüche 1-10.
40. Rekombinante DNA, dadurch gekennzeichnet, daß sie cDNA gemäß Anspruch 11 enthält.
45. Antigen, das unter Verwendung der cDNA gemäß Anspruch 11 oder der rekombinanten DNA gemäß Anspruch 12 hergestellt wurde oder unter Verwendung der Aminosäurenstruktur, die aus seiner cDNA abgeleitet werden kann.
50. Antigen nach Anspruch 13, dadurch gekennzeichnet, daß es ein Protein oder Peptid ist.
55. Antigen nach einem der Ansprüche 13 oder 14, dadurch gekennzeichnet, daß es eine Aminosäuresequenz aufweist, die der Tabelle 3 oder einer Teilsequenz davon entspricht.
60. Antigen nach Anspruch 15, dadurch gekennzeichnet, daß die Teilsequenz wenigstens 10 Aminosäuren aufweist.
65. Antigen nach Anspruch 15, dadurch gekennzeichnet, daß es die Aminosäuresequenz RLQALET-LIQNQQQLNLWGCKGKLICYTSVKWNTS oder eine Teilsequenz davon mit wenigstens 8 aufeinanderfolgenden Aminosäuren aufweist.
70. Antigen, das aus einem Immunschwäche-Virus gemäß einem der Ansprüche 1 bis 10 hergestellt wurde.
75. Antigen nach einem der Ansprüche 13 bis 18, dadurch gekennzeichnet, daß es rekombinant hergestellt wurde.

20. Antigen nach einem der Ansprüche 13 bis 17, dadurch gekennzeichnet, daß es synthetisch hergestellt wurde.
- 5 21. Testkit zum Nachweis von Antikörpern gegen Immunschwäche verursachende Viren, dadurch gekennzeichnet, daß Antigen gemäß den Ansprüchen 13 bis 20 eingesetzt wird.
- 10 22. Testkit gemäß Anspruch 21, dadurch gekennzeichnet, daß es ein Western Blot ist.
- 10 23. Testkit gemäß Anspruch 21, dadurch gekennzeichnet, daß es ein ELISA-Test ist oder ein Fluoreszenz-Antikörper-Nachweistest ist.
- 15 24. Verwendung des Immunschwäche-Virus gemäß einem der Ansprüche 1 bis 10 und/oder der cDNA gemäß Anspruch 11 oder 12 und/oder eines Antigens gemäß den Ansprüchen 13 bis 20 zum Nachweis von Retroviren, die Immunschwäche verursachen.
- 20 25. Verwendung eines Retrovirus nach einem der Ansprüche 1 bis 10, einer cDNA gemäß Anspruch 11 oder 12 und/oder eines Antigens gemäß den Ansprüchen 13 bis 20 zur Herstellung von Impfstoffen.
- 20 26. Ribonukleinsäure, dadurch gekennzeichnet, daß sie für ein Immunschwäche-Virus nach einem der Ansprüche 1 bis 10 kodiert.

26

30

36

40

45

50

55

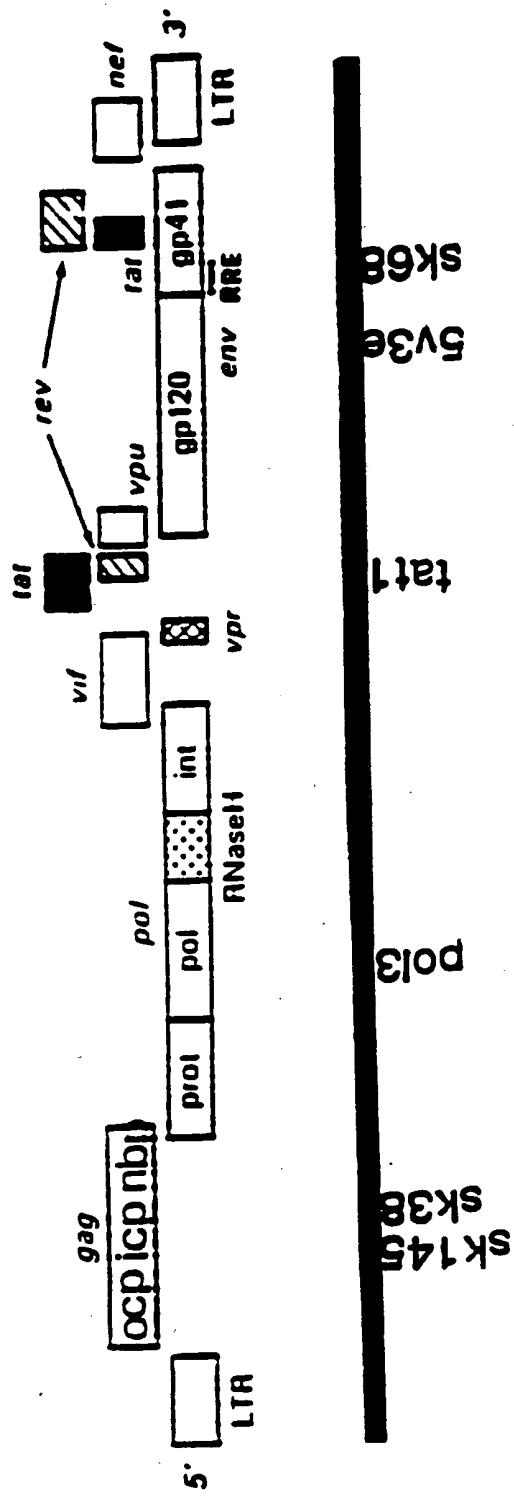


FIG. 1

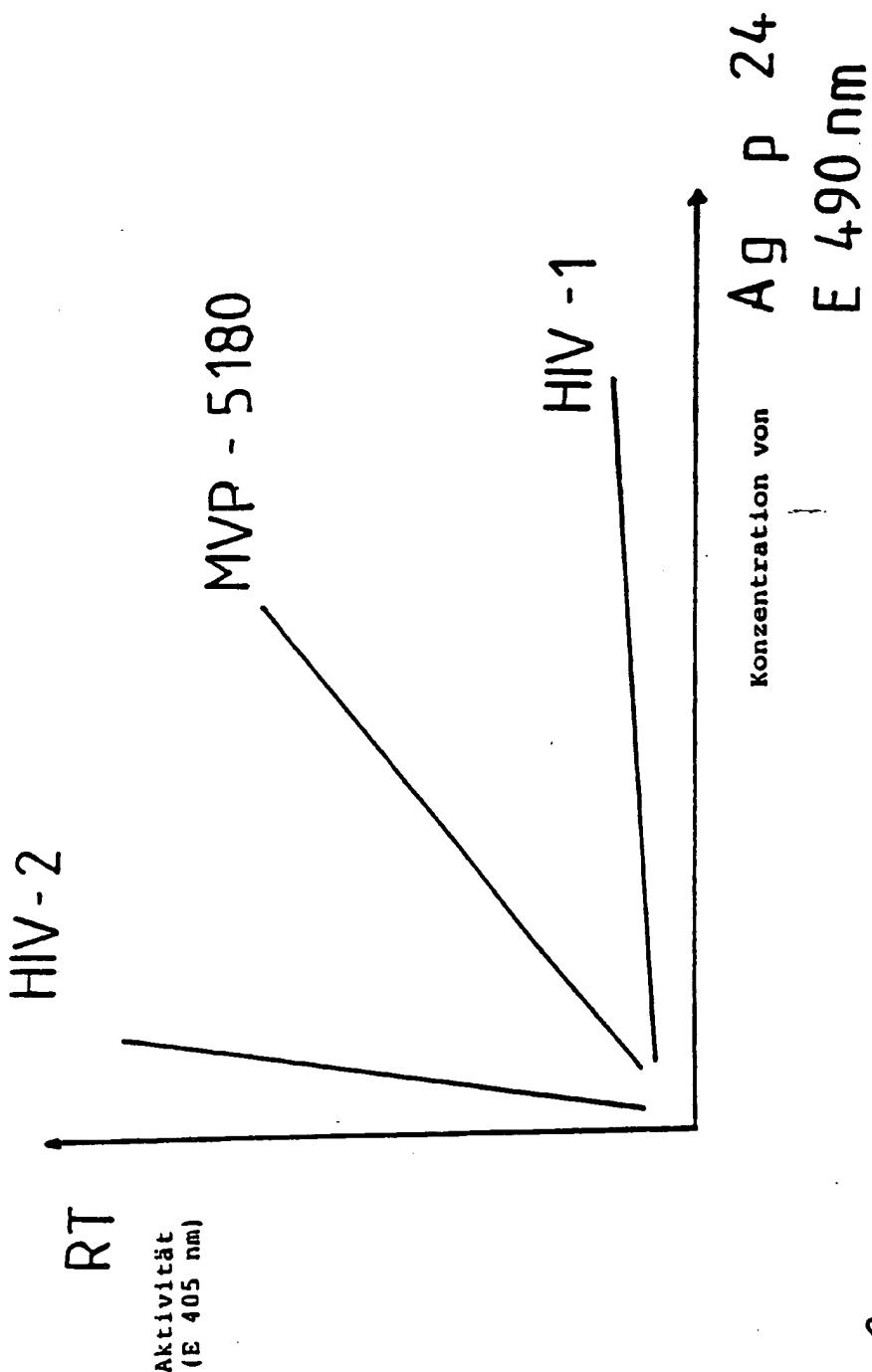


Fig. 2

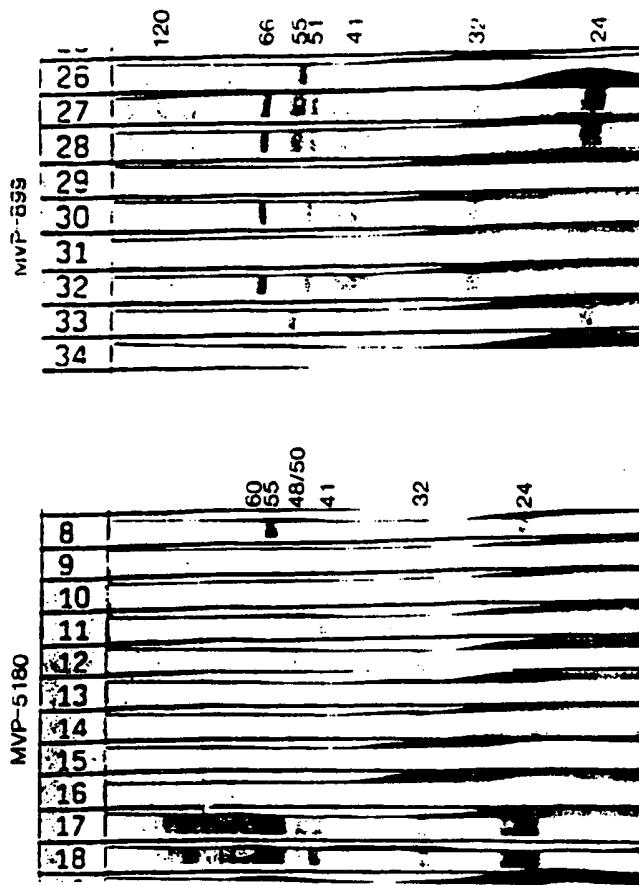


FIG. 3

Fig. 4: Sequenz von MVP 5180 (Sequ. ID No. 56)

```

1  CTGGATGGGT TAATTTACTC CCATAAGAGA GCAGAAATCC TGGATCTCTG
51  GATATATCAC ACTCAGGGAT TCTTCCCTGA TTGGCAGTGT TACACACCGG
101 GACCAGGACC TAGATTCCA CTGACATTTG GATGGTTGTT TAAACTGGTA
151 CCAGTGTCAg CAGAAGAGGC AGAGAGACTG GGTAATACAA ATGAAGATGC
201 TAGTCTTCTA CATCCAGCTT GTAATCATGG AGCTGAGGAT GCACACGGGG
251 AGATACTAAA ATGGCAGTTT GATAGATCAT TAGGCTTAAC ACATATAGCC
301 CTGCAAAAGC ACCCAGAGCT CTTCCCCAAG TAACTGACAC TGCAGGGACTT
351 TCCAGACTGC TGACACTGCG GGGACTTTCC AGCGTGGGAG GGATAAGGGG
401 CGGTTGGGG AGTGGCTAAC CCTCAGATGC TGCAATATAAG CAGCTGCTTT
451 CCGCTTGTAC CGGGTCTTAG TTAGAGGACC AGGTCTGAGC CCGGGAGCTC
501 CCTGGCCTCT AGCTGAACCC GCTGCTTAAC GCTCAATAAA GCTTGCCTTG
551 AGTGAGAAGC AGTGTGTGCT CATCTGTTCA ACCCTGGTGT CTAGAGATCC
601 CTCAGATCAC TTAGACTGAA GCAGAAAATC TCTAGCAGTG GCCCCCCAAC
651 AGGGACGCGA AAGTGAAAGT GGAACCAGGG AAGAAAACCT CCGACGCAAC
701 GGGCTCGGCT TAGCGGAGTG CACCTGCTAA GAGGCGAGAG GAACTCACAA
751 GAGGGTGAGT AAATTTGCTG GCGGTGGCCA GACCTAGGGG AAGGGCGAAG
801 TCCCTAGGGG AGGAAGATGG GTGCGAGAGC GTCTGTGTTG ACAGGGAGTA
851 AATTGGATGC ATGGGAACGA ATTAGGTTAA GGCCAGGATC TAAAAAGGCA
901 TATAGGCTAA AACATTTAGT ATGGGCAAGC AGGGACCTGG AAAGATAACGC
951 ATGTAATCCT GGTCTATTAG AAACGTGAGA AGGTACTGAG CAACTGCTAC
1001 AGCAGTTAGA GCCAGCTCTC AAGACAGGGT CAGAGGACCT GAAATCTCTC
1051 TGGAAACGCAA TAGCAGTACT CTGGTGCCTT CACAACAGAT TTGACATCCG
1101 AGATACACAG CAGGCAATAC AAAAGTTAAA GGAAGTAATG GCAAGCAGGA
1151 AGTCTGCAGA GGCCGCTAAG GAAGAAACAA GCCCTAGGCA GACAAGTCAGA
1201 AATTACCCCTA TAGTAACAAA TGCACAGGGA CAAATGGTAC ATCAAGCCAT

```

1251 CTCCCCCAGG ACTTTAAATG CATGGTAAA GGCAGTAGAA GAGAAGGCCT
 1301 TTAACCTGA AATTATTCCCT ATGTTATGG CATTATCAGA AGGGGCTGTC
 1351 CCCTATGATA TCAATACCAT GCTGAATGCC ATAGGGGGAC ACCAAGGGC
 1401 TTTACAAGTG TTGAAGGAAG TAATCAATGA GGAAGCAGCA GAATGGGATA
 1451 GAACTCATCC ACCAGCAATG GGGCCGTTAC CACCAGGGCA GATAAGGGAA
 1501 CCAACAGGAA GTGACATTGC TGGAACAACT AGCACACAGC AAGAGCAAAT
 1551 TATATGGACT ACTAGAGGGG CTAACCTAT CCCAGTAGGA GACATCTATA
 1601 GAAAATGGAT AGTGCTAGGA CTAAACAAAA TGGTAAAAAT GTACAGTCCA
 1651 GTGAGCATCT TAGATATTAG GCAGGGACCA AAAGAACCAT TCAGAGATTA
 1701 TGTAGATCGG TTTTACAAAA CATTAAAGAGC TGAGCAAGCT ACTCAAGAAG
 1751 TAAAGAATTG GATGACAGAA ACCTTGCTTG TTCAGAATTCA AAACCCAGAT
 1801 TGTAAACAAAA TTCTGAAAGC ATTAGGACCA GAAGCTACTT TAGAAGAAAT
 1851 GATGGTAGCC TGTCAAGGAG TAGGAGGGCC AACTCACAAG GCAAAAATAC
 1901 TAGCAGAACG AATGGCTTCT GCCCAGCAAG ATTTAAAAGG AGGATACACA
 1951 GCAGTATTCA TGCAAAGAGG GCAGAACATCA AATAGAAAAG GGCCCATAAA
 2001 ATGCTTCAAT TGTGGAAAAG AGGGACATAT AGCAAAAAC TGTGGAGCAC
 2051 CTAGAAAAAG GGGTTGCTGG AAATGTGGAC AGGAAGGTCA CCAAATGAAA
 2101 GATTGCAAAA ATGGAAGACA GGCAAATTTT TTAGGGAAGT ACTGGCCTCC
 2151 GGGGGGCACG AGGCCAGGCA ATTATGTGCA GAAACAAGTG TCCCCATCAG
 2201 CCCCCACCAAT GGAGGAGGCA GTGAAGGAAC AAGAGAACATCA GAGTCAGAAG
 2251 GGGGATCAGG AAGAGCTGTA CCCATTGCC TCCCTCAAAT CCCTCTTGG
 2301 GACAGACCAA TAGTCACAGC AAAGGTTGGG GGTCACTAT GTGAGGCTT
 2351 ACTGGATACA GGGCAGATG ATACAGTATT AAATAACATA CAATTAGAAG
 2401 GAAGATGGAC ACCAAAAATG ATAGGGGTA TAGGAGGCCTT TATAAAAGTA
 2451 AAAGAGTATA ACAATGTGAC AGTAGAAGTA CAAGGAAAGG AAGTACAGGG
 2501 AACAGTATTG GTGGGACCTA CTCCCTGTTAA TATTCTTGGG AGAAACATAT
 2551 TGACAGGATT AGGATGTACA CTAAATTCC CTATAAGTCC CATAGCCCCA

2601 GTGCCAGTAA AGCTAAAACC AGGAATGGAT GGACCAAAAG TAAAACAATG
 2651 GCCCCTATCT AGAGAGAAAA TAGAACACT AACTGCAATA TGTCAAGAAA
 2701 TGGAACAGGA AGGAAAATC TCAAGAATAG GACCTGAAAA TCCTTATAAT
 2751 ACACCTATTT TTGCTATAAA AAAGAAAGAT AGCACTAAGT GGAGAAAATT
 2801 GGTAGACTTC AGAGAATTAA ATAAAAGAAC ACAAGATTTC TGGGAGGTGC
 2851 AATTAGGTAT TCCACATCCA GGGGGTTAA AGCAAAGGCA ATCTGTTACA
 2901 GTCTTAGATG TAGGAGATGC TTATTCCTCA TGCCCTTTAG ATCCAGACTT
 2951 TAGAAAATAC ACTGCCTTCA CTATTCCTAG TGTGAACAAT GAGACCCCAG
 3001 GAGTAAGATA CCAGTACAAT GTCCTCCGC AAGGGTGGAA AGGTTCACCA
 3051 GCCATATTC AGAGTTCAAT GACAAAGATT CTAGATCCAT TTAGAAAAG
 3101 CAACCCAGAA GTAGAAATTG ATCAGTACAT AGATGACTTA TATCTAGGAT
 3151 CAGATTTACC ATTGGCAGAA CATAGAAAGA GGGTCGAATT GCTTAGGGAA
 3201 CATTATATC AGTGGGGATT TACTACCCCT GATAAAAAGC ATCAGAAGGA
 3251 ACCTCCCTTT TTATGGATGG GATATGAGCT CCACCCAGAC AAGTGGACAG
 3301 TACAGCCCAT CCAATTGCCT GACAAAGAAG TGTGGACAGT AAATGATATA
 3351 CAAAAATTAG TAGGAAAATT AAATTGGCA AGTCAAATCT ATCAAGGAAT
 3401 TAGAGTAAAA GAATTGTGCA AGTTAATCAG AGGAACCAAA TCATTGACAG
 3451 AGGTAGTACC TTTAAGTAAA GAGGCAGAAC TAGAATTAGA AGAAAACAGA
 3501 GAAAAGCTAA AAGAGCCAGT ACATGGAGTA TATTACCGAC CTGACAAAGA
 3551 CTTGTGGGTT AGTATTCAAG AGCATGGAGA AGGGCAATGG ACTTACCAAGG
 3601 TATATCAGGA TGAACATAAG AACCTAAAAA CAGGAAAATA TGCTAGGCAA
 3651 AAGGCCTCCC ACACAAATGA TATAAGACAA TTGGCAGAAC TAGTCCAGAA
 3701 GGTGTCTCAA GAAGCTATAG TTATATGGGG GAAATTACCT AAATTCAGGC
 3751 TGCCAGTTAC TAGAGAAACT TGGGAAACTT GGTGGGCAGA ATATTGGCAG
 3801 GCCACCTGGA TTCCTGAATG GGAATTGTC AGCACACCCC CATTGATCAA
 3851 ATTATGGTAC CAGTTAGAAA CAGAACCTAT TGTAGGGGCA GAAACCTTTT
 3901 ATGTAGATGG AGCAGCTAAT AGGAATACAA AACTAGGAAA GGCGGGATAT

3951 GTTACAGAAC AAGGAAAACA GAACATAATA AAGTTAGAAG AGACAACCAA
4001 TCAAAAGGCT GAATTAATGG CTGTATTAAT AGCCTTGCAG GATTCCAAGG
4051 AGCAAGTAAA CATACTAACCA GACTCACAAT ATGTATTGGG CATCATATCC
4101 TCCCCAACCAA CACAGAGTGA CTCCCCATAA GTTCAGCAGA TAATAGAGGA
4151 ACTAACAAAAA AAGGAACGAG TGTATCTTAC ATGGGTTCCCT GCTCACAAAG
4201 GCATAGGAGG AAATGAAAAA ATAGATAAT TAGTAAGCAA AGACATTAGA
4251 AGAGTCCTGT TCCTGGAAGG AATAGATCAG GCACAAGAAG ATCATGAAAA
4301 ATATCATAGT AATTGGAGAG CATTAGCTAG TGACTTTGGA TTACCACCAA
4351 TAGTAGCCAA GGAAATCATT GCTAGTTGTC CTAATGCCA TATAAAAGGG
4401 GAAGCAACGC ATGGTCAAGT AGACTACAGC CCAGAGATAT GGCAAATGGA
4451 TTGTACACAT TTAGAAGGCA AAATCATAAT AGTTGCTGTC CATGTAGCAA
4501 GTGACTTTAT AGAAGCAGAG GTGATAACAG CAGAAACAGG ACAGGAAACT
4551 GCCTATTTCC TGTTAAAATT AGCAGCAAGA TGGCCTGTCA AAGTAATACA
4601 TACAGACAAT GGACCTAATT TTACAAGTGC AGCCATGAAA GCTGCATGTT
4651 GGTGGACAGG CATAAACAT GAGTTGGGA TACCATATAA TCCACAAAGT
4701 CAAGGAGTAG TAGAAGCCAT GAATAAGAA TTAAAATCTA TTATACAGCA
4751 GGTGAGGGAC CAAGCAGAGC ATTTAAAAC AGCAGTACAA ATGGCAGTCT
4801 TTGTTCACAA TTTTAAAAGA AAAGGGGGGA TTGGGGGTA CACTGCAGGG
4851 GAGAGACTAA TAGACATACT AGCATCACAA ATACAAACAA CAGAACTACA
4901 AAAACAAATT TTAAAAATCA ACAATTTCG GGTCTATTAC AGAGATAGCA
4951 GAGACCCTAT TTGGAAAGGA CCGGCACAAAC TCCTGTGGAA AGGTGAGGG
5001 GCAGTAGTCA TACAAGATAA AGGAGACATT AAAGTGGTAC CAAGAAGAAA
5051 GGCAAAATA ATCAGAGATT ATGGAAAACA GATGGCAGGT ACTGATAGTA
5101 TGGCAAATAG ACAGACAGAA AGTGAAGCA TGGAACAGCC TGGTGAATA
5151 CCATAAAATAC ATGTCTAAGA AGGCCCGCAA CTGGCGTTAT AGGCATCATT
5201 ATGAATCCAG GAATCCAAA GTCAGTTCGG CGGTGTATAT TCCAGTAGCA
5251 GAAGCTGATA TAGTGGTCAC CACATATTGG GGATTAATGC CAGGGAAAG

5301 AGAGGAACAC TTGGGACATG GGGTTAGTAT AGAATGGCAA TACAAGGAGT
5351 ATAAAACACA GATTGATCCT GAAACAGCAG ACAGGATGAT ACATCTGCAT
5401 TATTTCACAT GTTTACAGA ATCAGCAATC AGGAAGGCCA TTCTAGGGCA
5451 GAGAGTGCTG ACCAAGTGTG AATACCTGGC AGGACATAGT CAGGTAGGGA
5501 CACTACAATT CTTAGCCTTG AAAGCAGTAG TGAAAGTAAA AAGAAATAAG
5551 CCTCCCCCTAC CCAGTGTCCA GAGATTAACA GAAGATAGAT GGAACAAGCC
5601 CTGGAAAATC AGGGACCAGC TAGGGAGCCA TTCAATGAAT GGACACTAGA
5651 GCTCCTGGAA GAGCTGAAAG AAGAACAGT AAGACATTTC CCTAGGCCTT
5701 GGTTACAAGC CTGTGGGCAG TACATTTATG AGACTTATGG AGACACTTGG
5751 GAAGGAGTTA TGGCAATTAT AAGAATCTTA CAACAACTAC TGTTTACCCA
5801 TTATAGAATT GGATGCCAAC ATAGTAGAAT AGGAATTCTC CCATCTAACCA
5851 CAAGAGGAAG AGGAAGAAGA AATGGATCCA GTAGATCCTG AGATGCC
5901 TTGGCATCAC CCTGGGAGCA AGCCCCAAC CCCTTGTAAAT AATTGCTATT
5951 GCAAAAGATG CTGCTATCAT TGCTATGTT GTTTCACAAA GAAGGGTTG
6001 GGAATCTCCC ATGGCAGGAA GAAGCGAAGA AGACCAGCAG CTGCTGCAAG
6051 CTATCCAGAT AATAAAGATC CTGTACCAAGA GCAGTAAGTA ACGCTGATGC
6101 ATCAAGAGAA CCTGCTAGCC TTAATAGCTT TAAGTGCTTT GTGTCTTATA
6151 AATGTACTTA TATGGTTGTT TAACCTTAGA ATTTATTAG TGCAAAGAAA
6201 ACAAGATAGA AGGGAGCAGG AAATACTTGA AAGATTAAGG AGAATAAAGG
6251 AAATCAGGGA TGACAGTGAC TATGAAAGTA ATGAAGAAGA ACAACAGGAA
6301 GTCATGGAGC TTATACATAG CCATGGCTTT GCTAATCCCA TGTTTGAGTT
6351 ATAGTAAACA ATTGTATGCC ACAGTTTATT CTGGGGTACC TGTATGGAA
6401 GAGGCAGCAC CAGTACTATT CTGTGCTTCA GATGCTAACCC TAACAAGCAC
6451 TGAACAGCAT AATATTTGGG CATCACAAAGC CTGCGTTCCCT ACAGATCCCA
6501 ATCCACATGA ATTTCCACTA GGCAATGTGA CAGATAACTT TGATATATGG
6551 AAAAATTACA TGGTGGACCA AATGCATGAA GACATCATTA GTTGTGGGA
6601 ACAGAGTTA AAGCCTTGTG AGAAAATGAC TTTCTTATGT GTACAAATGA

6651 ACTGTGTTAGA TCTGCAAACA AATAAAACAG GCCTATTAAA TGAGACAATA
 6701 AATGAGATGA GAAATTGTAG TTTTAATGTA ACTACAGTCC TCACAGACAA
 6751 AAAGGAGCAA AAACAGGCTC TATTCTATGT ATCAGATCTG AGTAAGGTTA
 6801 ATGACTCAAA TGCAGTAAAT GGAACAACAT ATATGTTAAC TAATTGTAAC
 6851 TCCACAATTAA TCAAGCAGGC CTGTCCGAAG GTAAGTTTG AGCCCATTCC
 6901 CATAACTAT TGTGCTCCAA CAGGATATGC CATCTTTAAC TGTAATGACA
 6951 CAGACTTTAA TGGAACAGGC CTATGCCACA ATATTCAGT GGTTACTTGT
 7001 ACACATGGCA TCAAGCCAAC AGTAAGTACT CAACTAATAC TGAATGGAC
 7051 ACTCTCTAGA GAAAAGATAA GAATTATGGG AAAAAATATT ACAGAACATCAG
 7101 CAAAGAATAT CATACTAACCC CAAACACTC CTATAAACAT GACCTGCATA
 7151 AGAGAAGGAA TTGCAGAGGT ACAAGATATA TATACAGGTC CAATGAGATG
 7201 GCGCAGTATG ACACTTAAAAA GAAGTAACAA TACATCACCA AGATCAAGGG
 7251 TAGCTTATTG TACATATAAT AAGACTGTAT GGGAAATGC CCTACAACAA
 7301 ACAGCTATAA GGTATTTAAA TCTTGTAAAC CAAACAGAGA ATGTTACCAT
 7351 AATATTCAAGC AGAACTAGTG GTGGAGATGC AGAAGTAAGC CATTACATT
 7401 TTAACTGTCA TGGAGAATTG TTTTATTGTA ACACATCTGG GATGTTAAC
 7451 TATACTTTA TCAACTGTAC AAAGTCCGGA TGCCAGGAGA TCAAAGGGAG
 7501 CAATGAGACC AATAAAATG GTACTATACC TTGCAAGTTA AGACAGCTAG
 7551 TAAGATCATG GATGAAGGGA GAGTCGAGAA TCTATGCACC TCCCATCCCC
 7601 GGCAACTTAA CATGTCATTC CAACATAACT GGAATGATTG TACAGTTAGA
 7651 TCAACCATGG AATTCCACAG GTGAAAATAC ACTTAGACCA GTAGGGGGAG
 7701 ATATGAAAGA TATATGGAGA ACTAAATTGT ACAACTACAA AGTAGTACAG
 7751 ATAAAACCTT TTAGTGTAGC ACCTACAAAA ATGTCAAGAC CAATAATAAA
 7801 CATTACACACC CCTCACAGGG AAAAAAGAGC AGTAGGATTG GGAATGCTAT
 7851 TCTTGGGGGT GCTAAGTGCA GCAGGTAGCA CTATGGGCGC AGCGGCAACA
 7901 GCGCTGACGG TACGGACCCA CAGTGTACTG AAGGGTATAAG TGCAACAGCA
 7951 GGACAACTTG CTGAGAGCGA TACAGGCCA GCAACACTTG CTGAGGTTAT

8001 CTGTATGGGG TATTAGACAA CTCCGAGCTC GCCTGCAAGC CTTAGAAACC
 8051 CTTATACAGA ATCAGCAACG CCTAACCTA TGGGGCTGTA AAGGAAACT
 8101 AATCTGTTAC ACATCAGTAA AATGGAACAC ATCATGGTCA GGAAGATATA
 8151 ATGATGACAG TATTTGGGAC AACCTTACAT GGCAGCAATG GGACCAACAC
 8201 ATAAACAATG TAAGCTCCAT TATATATGAT GAAATACAAG CAGCACAAAGA
 8251 CCAACAGGAA AAGAATGTA AAGCATTGTT GGAGCTAGAT GAATGGGCCT
 8301 CTCTTGGAA TTGGTTGAC ATAACATAAT GGTTGTGTA TATAAAAATA
 8351 GCTATAATCA TAGTGGGAGC ACTAATAGGT ATAAGAGTTA TTATGATAAT
 8401 ACTTAATCTA GTGAAGAACAA TTAGGCAGGG ATATCAACCC CTCTCGTTGC
 8451 AGATCCCTGT CCCACACCGG CAGGAAGCAG AAACGCCAGG AAGAACAGGA
 8501 GAAGAAGGTG GAGAAGGAGA CAGGCCAAG TGGACAGCCT TGCCACCAGG
 8551 ATTCTTGCAA CAGTTGTACA CGGATCTCAG GACAATAATC TTGTGGACTT
 8601 ACCACCTCTT GAGCAACTTA ATATCAGGG A TCCGGAGGCT GATCGACTAC
 8651 CTGGGACTGG GACTGTGGAT CCTGGGACAA AAGACAATTG AAGCTTGTAG
 8701 ACTTTGTGGA GCTGTAATGC AATATTGGCT ACAAGAATTG AAAAATAGTG
 8751 CTACAAACCT GCTTGATACT ATTGCAGTGT CAGTTGCCAA TTGGACTGAC
 8801 GGCATCATCT TAGGTCTACA AAGAATAGGA CAAGGATTCC TTCACATCCC
 8851 AAGAAGAATT AGACAAGGTG CAGAAAGAAT CTTAGTGTAA CATGGGAAT
 8901 GCATGGAGCA AAAGCAAATT TGCAGGATGG TCAGAAGTAA GAGATAGAAT
 8951 GAGACGATCC TCCTCTGATC CTCAACAACC ATGTGCACCT GGAGTAGGAG
 9001 CTGTCTCCAG GGAGTTAGCA ACTAGAGGGG GAATATCAAG TTCCCACACT
 9051 CCTCAAAACA ATGCAGCCCT TGCATTCTA GACAGCCACA AAGATGAGGA
 9101 TGTAGGCTTC CCAGTAAGAC CTCAAGTGCC TCTAAGGCCA ATGACCTTTA
 9151 AAGCAGCCTT TGACCTCAGC TTCTTTAA AAGAAAAGGG AGGACTGGAT
 9201 GGGTTAATT ACTCCCATAA GAGAGCAGAA ATCCTGGATC TCTGGATATA
 9251 TCACACTCAG GGATTCTTCC CTGATTGGCA GTGTTACACA CCGGGACCAG
 9301 GACCTAGATT CCCACTGACA TTTGGATGGT TGTTAAACT GGTACCAGTG

9351 TCAGCAGAAG AGGCAGAGAG ACTGGGTAAT ACAAAATGAAG ATGCTAGTCT
9401 TCTACATCCA GCTTGTAATC ATGGAGCTGA GGATGCACAC GGGGAGATAAC
9451 TAAAATGGCA GTTTGATAGA TCATTAGGCT TAACACATAT AGCCCTGCAA
9501 AAGCACCCAG AGCTCTTCCC CAAGTAACTG AACTGCGGG ACTTTCCAGA
9551 CTGCTGACAC TCGGGGGACT TTCCAGCGTG GGAGGGATAAA GGGGCGGTTC
9601 GGGGAGTGGC TAACCCTCAG ATGCTGCATA TAAGCAGCTG CTTTCCGCTT
9651 GTACCGGGTC TTAGTTAGAG GACCAGGTCT GAGCCCGGGGA GCTCCCTGGC
9701 CTCTAGCTGA ACCCGCTGCT TAACGCTCAA TAAAGCTTGC CTTGAGTGAG
9751 AAGCAGTGTG TGCTCATCTG TTCAACCCTG GTGTCTAGAG ATC

Figur 5: PCR Amplifizierungs-, Klonierungs- und Sequenzierungsstrategie:

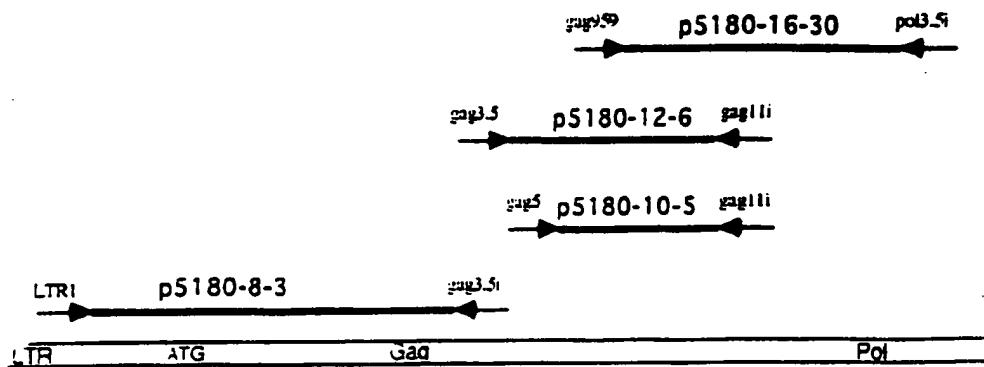


Fig. 6: obere Zeile entspricht Abb. 4, untere Zeile mit PCR-Technik ermittelt (Sequ. No. 57, 58)

MvP5180	685	AAACCTCCGACGCAACGGGCTCGGCTTAGCGGAGTGCACCTGCTAAGAGG	734
	1	aaacctccaacgcaacggctcgcttagcgagtgacacgtctaagagg	50
	735	CGAGAGGAACTCACAAGAGGGTGAGTAAATTGCTGGCGGTGCCAGACC	784
	51	cgagaggaactcacaagagggtgagtaaattgctggcggtggccagacc	100
	785	TAGGGGAAGGGCGAAGTCCCTAGGGAGGAAGATGGGTGCGAGAGCGTCT	834
	101	taggggaaggggcgaagtccctagggaggaagatgggtgcgagacggct	150
	835	GTGTTGACAGGGAGTAAATTGGATGCATGGAACGAATTAGGTTAAGGCC	884
	151	gtgttgacagggagtaaattggatgcattggaaacaaatttaggttatggcc	200
	885	AGGATCTAAAAGGCATATAGGCTAAACATTTAGTATGGCAAGCAGGG	934
	201	aggatctaaaaggcatataggctaaacatttatgttatggcaagcaggg	250
	935	AGCTGGAAAGATAACGCATGTAATCCTGGTCTATTAGAAACTGCAGAAGGT	984
	251	agctggaaagataacgcataatcctggtctactagaaaactgcagaaggt	300
	985	ACTGAGCAACTGCTACAGCAGTTAGGCCAGCTCTCAAGACAGGGTCAGA	1034
	301	actgacaaactgtacagcagtttagccagctctcaagacagggcaga	350
	1035	GGACCTGAAATCTCTCTGGAACGCAATAGCAGTACTCTGGTCCGTTCA	1084
	351	ggacctgaaatccctctggaaacgcaatagcagttactctggtgcgttcaca	400
	1085	ACAGATTTGACATCCGAGATAACACAGCAGGCCAATACAAAAGTTAAGGAA	1134
	401	acagatggacatccgagataacacagcaggcaatacaatggaa	450
	1135	GTAATGGCAAGCAGGAAGTCTGCAGAGGGCCGCTAAGGAAGAAACAGCCC	1184
	451	gtaatggcaagcaggaagtctgcagaggccgctaaggaaagaacaagctc	500

1185 TAGGCAGACAAGTCAAAATTACCTATAGTAACAAATGCACAGGGACAAA 1234
 501 aaggcaggcaagtcaaaattaccctatacgtaacaaatgcacagggacaaa 550

1235 TGGTACATCAAGCCATCTCCCCAGGACTTTAAATGCATGGGTAAAGGCA 1284
 551 tggtacatcaagccatatacccttaggactttaatgcatgggtaaaggca 600

1285 GTAGAAGAGAAGGCCTTTAACCTGAAATTATTCCCTATGTTATGGCATT 1334
 601 gtagaagaaaaggccttaaccctgaaattattccctatgttatggcatt 650

1335 ATCAGAAGGGCTGTCCCCATGATATCAATACCAGCTGAATGCCATAG 1384
 651 atcagaaggggctgtccctatgatatacataaccatgctgaatgccatag 700

1385 GGGGACACCAAGGGGCTTTACAAGTGTGAAGGAAGTAATCAATGAGGAA 1434
 701 ggggacaccaaggggcttacaagtgtgaaggaagtaatcaatgaggaa 750

1435 GCAGCAGAATGGGATAGAACTCATCCACCAAGCAATGGGGCCGTTACCACC 1484
 751 gcagcagattgggatagaactcatccaccagcaatggggccgttaccacc 800

1485 AGGGCAGATAAGGGAACCAAACAGGAAGTGACATTGCTGGAAACAACTAGCA 1534
 801 agggcagataagggaaccaacaggaagtgacattgctggaaacaactagca 850

1535 CACAGCAAGAGCAAATTATGGACTACTAGAGGGGCTAACTCTATCCCA 1584
 851 cacagcaagagcaaattatggactactagagggctaaactctatccca 900

1585 GTAGGAGACATCTATAGAAAATGGATAGTGCTAGGACTAAACAAAATGGT 1634
 901 gtaggagacatctatagaaaatggatagtgttaggactaaacaaaatgg 950

1635 AAAAATGTACAGTCCAGTGAGCATCTTAGATATTAGGCAGGGACAAAG 1684
 951 aaaaatgtacagtccagtgagcatcttagatattaggcagggacaaaag 1000

1685	AACCATTCAAGAGATTATGTAGATCGGTTTACAAAACATTAAGAGCTGAG	1734
1001	aaccattcaagagattatgttagatcggtttacaaaacattaagagctgag	1050
1735	CAAGCTACTCAAGAAGTAAAGAATTGGATGACAGAACCTTGCTTGTCA	1784
1051	caagctactcaagaagtaaaagaattggatgacagaaaaccctcggttca	1100
1785	GAATTCAAACCCAGATTGTAAACAAATTCTGAAAGCATTAGGACCAGAAG	1834
1101	gaattcaaaccagattgtaaacaaattctgaaagcattaggaccaggag	1150
1835	CTACTTTAGAAGAAATGATGGTAGCCTGTCAAGGAGTAGGGAGGCCAACT	1884
1151	ctactttagaagaaatgatggtagcctgtcaaggagtagggaggccaaact	1200
1885	CACAAGGCAGAAATACTAGCAGAAGCAATGGCTCTGCCAGCAAGATT	1934
1201	cacaaggcaaaaatactagcagaagcaatggctctgcccagcaagattt	1250
1935	AAAAGGAGGATACACAGCAGTATTCATGCAAAGAGGGCAGAATCCAAATA	1984
1251	aaagggaggatacacagcagtttcatgaaagaggcagaatccaaata	1300
1985	GAAAAGGGCCATAAAATGCTCAATTGTGGAAAGAGGGACATATAGCA	2034
1301	aaaaaggccctataaaatgttcaattgtggaaaagaggacatatacgca	1350
2035	AAAAACTGTGAGCACCTAGAAAAAGGGGTTGCTGGAAATGTGGACAGGA	2084
1351	aaaaactgtcgagcacctagaagaagggttactggaaatgtggacagga	1400
2085	AGGTCAACCAATGAAAGATTGCAAAAATGGAAGACAGGCCAAATTTTTAG	2134
1401	aggtcaccaaattgaaagattgcaaaaatggaagacaggctatttttag	1450
2135	GGAAAGTACTGGCCTCCGGGGGGCACGAGGCCAGGCAATTATGTGCAGAAA	2184
1451	ggaagtactggcctccggggggcacgaggccagccaattatgtgcagaaa	1500

2185	CAAGTGTCCCCATCAGCCCCACCAATGGAGGAGGCAGTGAAGGAACAAGA	2234
1501	caagtgtccccatcagccccaccaatggaggaggcagtgaaggacaaga	1550
2235	GAATCAGAGTCAGAAGGGGGATCAGGAAGAGCTGTACCCATTGCCCTCCC	2284
1551	aatcagaatcaaaagggggatcaggaagagctgtacccattgcctccc	1600
2285	TCAAATCCCTCTTGGGACAGACCAATAGTCACAGCAAAGGTTGGGGTC	2334
1601	tcaaatccctcttggacagaccaatagtcacagcaaagggttggggcc	1650
2335	ATCTATGTGAGGCTTACTGGATACAGGGGCAGATGATACTAGTATTAAAT	2384
1651	atctatgtgaggcttactggatacaggggcagatgatacagtattaaat	1700
2385	AACATACAATTAGAAGGAAGATGGACACCAAAA	2417
1701	aacatacaattagaagatggacacccaaa	1733

Fig. 7: Vergleich des gag-Proteins, einmal gemäß Abb. 1 (jeweils oben) und einmal mit PCR-Technik (jeweils unten) ermittelt (Sequ. ID No. 56, 60)

MGARASVLTGSKLDAYERIRLRPGSKKAYRLKHLVWASRELERYACNPGL
MGARRSVLTGSKLDAYERIRLRPGSKKAYRLKHLVWASRELERYAYNPGL

LETAEGTEQLLQQLEPALKTGSDELKSLWNIAAVLWCVHNRFDIRDTQQA
LETAEGTEQLLQQLEPALKTGSDELKSLWNIAAVLWCVHNRFDIRDTQQA

IQKLKEVMASRKSAEAAKEETSPRQTSQNYPIVTNAQGQMVKQAIISPRTL
IQKLKEVMASRKSAEAAKEETSSSTQASQNYPIVTNAQGQMVKQAIISPRTL

NAWVKAVEEKAFNPEIIPMFMASEGAVPYDINTMLNAIGGHQGALQVLK
NAWVKAVEEKAFNPEIIPMFMASEGAVPYDINTMLNAIGGHQGALQVLK

EVINEEEAAEWDRTHPAMGPLPPGQIREPTGSDIAGTTSTQQEQIIWTTR
EVINEEEAADWDRTHPAMGPLPPGQIREPTGSDIAGTTSTQQEQIIWTTR

GANSIPVGDIYRKWIVLGLNKMVKMYSPPSILDIRQGPKEFPRDYVDRFY
GANSIPVGDIYRKWIVLGLNKMVKMYSPPSILDIRQGPKEFPRDYVDRFY

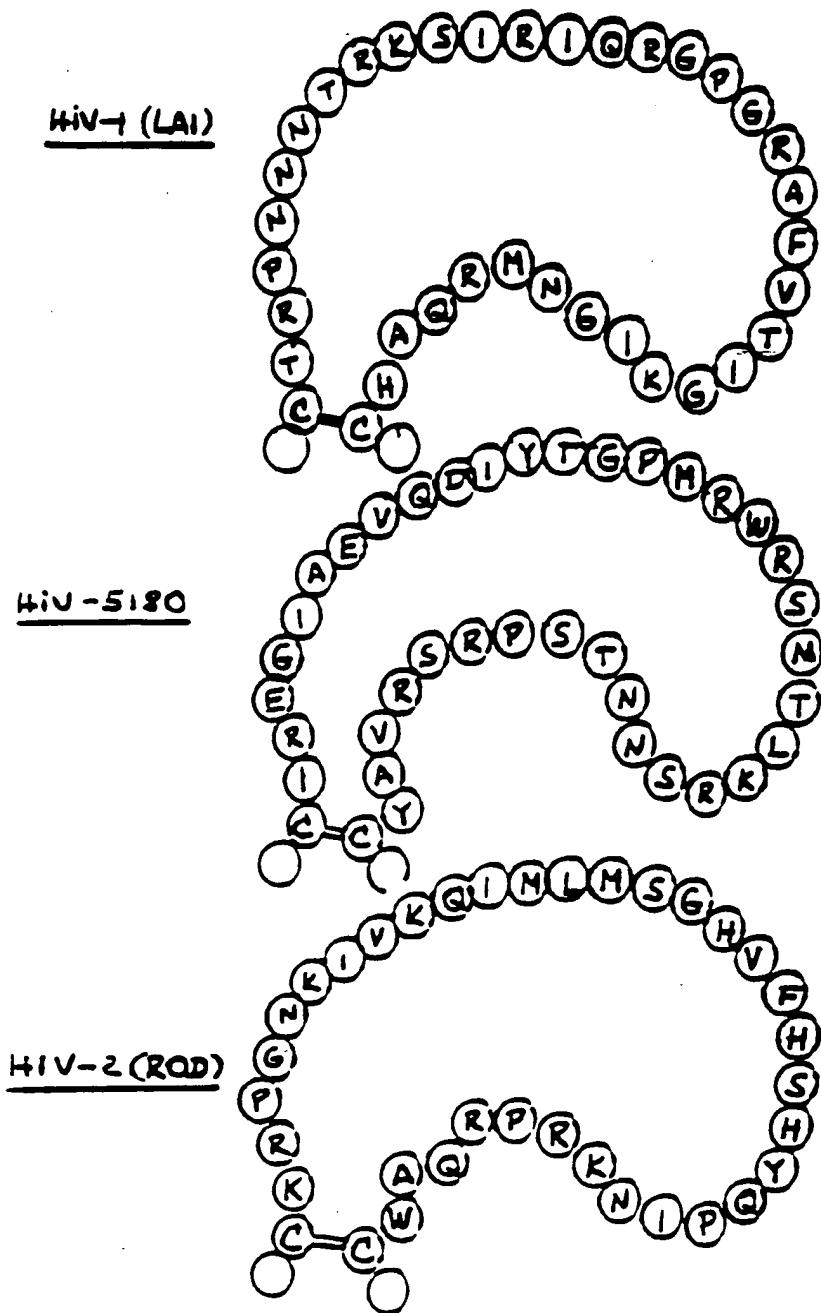
KTLRAEQATQEVKNWMTETLLVQNSNPCKQILKALGPEATLEEMMVACQ
KTLRAEQATQEVKNWMTETLVVQNSNPCKQILKALGPGATLEEMMVACQ

GVGGPTHKAKILAEAMASAQQDLKGGYTAVFMQRGQNPNRKGPIKCFNCG
GVGGPTHKAKILAEAMASAQQDLKGGYTAVFMQRGQNPNRKGPIKCFNCG

KEGHIAKNCRAPRKGCWKCGQEGHQMKDCNGRQANFLGKYWPPGGTRP
KEGHIAKNCRAPRRGYWKCGQEGHQMKDCNGRQANFLGKYWPPGGTRP

GNVYQKQVSPSAPPMEAVKEQENQSQKGDQEELYPPFASLKSLEFGTDQ
ANYVQKQVSPSAPPMEAVKEQENQNQKGDQEELYPPFASLKSLEFGTDQ

Fig. 8



THIS PAGE BLANK (USPTO)